"A Stochastic Model of Space-Time Variability of Tropical Rainfall: I. Statistics of Spatial Averages"

Prasun K. Kundu[1] and Thomas L. Bell[2]

Submitted to Water Resources Research

Popular Summary

Global maps of rainfall are of great importance in connection with modeling of the earth's climate. Comparison between the maps of rainfall predicted by computer-generated climate models with observation provides a sensitive test for these models. To make such a comparison, one typically needs the total precipitation amount over a large area, which could be hundreds of kilometers in size over extended periods of time of order days or months. This presents a difficult problem since rain varies greatly from place to place as well as in time.

Remote sensing methods using ground radar or satellites detect rain over a large area by essentially taking a series of snapshots at infrequent intervals and indirectly deriving the average rain intensity within a collection of "pixels", usually several kilometers in size. They measure area average of rain at a particular instant. Rain gauges, on the other hand, record rain accumulation continuously in time but only over a very small area tens of centimeters across, say, the size of a dinner plate. They measure only a time average at a single location. In making use of either method one needs to fill in the gaps in the observation – either the gaps in the area covered or the gaps in time of observation. This involves using statistical models to obtain information about the rain that is missed from what is actually detected. This paper investigates such a statistical model and validates it with rain data collected over the tropical Western Pacific from ship borne radars during TOGA COARE (Tropical Oceans Global Atmosphere Coupled Ocean-Atmosphere Response Experiment).

The model incorporates a number of commonly observed features of rain. While rain varies rapidly with location and time, the variability diminishes when averaged over larger areas or longer periods of time. Moreover, rain is patchy in nature – at any instant on the average only a certain fraction of the observed pixels contain rain. The fraction of area covered by rain decreases, as the size of a pixel becomes smaller. This means that within what looks like a patch of rainy area in a coarse resolution view with larger pixel size, one finds clusters of rainy and dry patches when viewed on a finer scale. The model makes definite predictions about how these and other related statistics depend on the pixel size. These predictions were found to agree well with data. In a subsequent second part of the work we plan to test the model with rain gauge data collected during the TRMM (Tropical Rainfall Measuring Mission) ground validation campaign.

[1]GEST/UMBC
[2]NASA/GSFC

Submitted to *Water Resources Res.* October, 2002

A Stochastic Model of Space-Time Variability of Tropical Rainfall: I. Statistics of Spatial Averages

Prasun K. Kundu

Goddard Earth Sciences and Technology Center, University of Maryland Baltimore County, Baltimore, MD and NASA/Goddard Space Flight Center, Greenbelt, MD

Thomas L. Bell

Laboratory for Atmospheres, NASA/Goddard Space Flight Center, Greenbelt, MD

Abstract. A characteristic feature of rainfall statistics is that they depend on the space and time scales over which rain data are averaged. As part of an earlier effort to determine the sampling error of satellite rain averages, a space-time model of rain statistics was developed to describe the statistics of gridded rain observed over the eastern tropical Atlantic. The model allows one to compute the second moment statistics of space- and time-averaged rain rate, which can be fitted to satellite or rain gauge data to determine the four model parameters – an overall strength parameter, a characteristic length separating the long and short wavelength regimes, a characteristic relaxation time for decay of the autocorrelation of the instantaneous local rain rate, and a certain "fractal" power law exponent. For area-averaged instantaneous rain rate, this exponent governs the power law dependence of these statistics on the averaging length scale L. In particular, the variance of rain rate averaged over an $L \times L$ area exhibits a power law singularity as $L \to 0$. In the present work a more efficient method of estimating the model parameters is presented, and used to fit the model to the statistics of area-averaged rain rate over the tropical Western Pacific measured with ship-borne radar during TOGA COARE. Good agreement is found between the data and predictions from the model over a wide range of averaging scales. An extension of the spectral model scaling relations to describe the dependence of the fractional coverage of rain on the spatial scale L is also explored.

1. Introduction

Rainfall is a complex phenomenon involving the interplay of many physical processes in the atmosphere. This makes it extremely difficult to quantitatively describe rainfall in terms of a detailed model of the underlying processes. Instead, a mathematically tractable and yet realistic description naturally calls for statistical considerations. In contrast with the physical models, models of rain statistics have the advantage of being conceptually economical in the sense that they generally involve fewer adjustable parameters. Moreover, they can be easily validated by comparing against real data gathered over a large area and a large time period representative of a certain climatological regime. Once the model parameters are determined from a sufficiently large data set, the model provides an efficient method of describing various statistical properties of rainfall over areas with similar rain climatologies [see, e.g., *Bell et al.*, 2001, hereafter referred to as BKK]. For many purposes, this is all that is needed.

It is common experience that rainfall is highly irregular in space and time and difficult to accurately predict and measure. Rain gauges have traditionally

provided a direct measurement of rain on a local basis. However, the inherently sporadic nature of rainfall along with the practical difficulty of laying down a sufficiently dense network of continuously monitored gauges makes this method ill-suited to the problem of measuring precipitation on a global scale over an extended time period. In recent years, remote sensing techniques have provided an alternative and more convenient method of rain estimation. These techniques are attractive because of their ability to cover large areas at a time. Remote sensing measurements, when carried out from low earth-orbiting satellites, enable one to estimate rainfall over a truly global scale.

Inherent in these remote sensing measurements is a finite spatial resolution, since these methods infer the rain rate indirectly from the intensity of the radiation that carries information about precipitation present within a certain volume of the atmosphere. This resolution is governed by the spatial resolution of the instrument's antennae. Thus, the primary quantity estimated by these methods from the retrieval output is the area-average of the local instantaneous rain rate $R(\mathbf{x},t)$ over a spatial domain $\mathcal{A}$ of area A , which approximates the effective area covered by the instrument (whose actual response profile may be quite complex),

$$\mathcal{R}_{\mathcal{A}}(t) = \frac{1}{A}\int_{\mathcal{A}} d^2\mathbf{x}\, R(\mathbf{x},t)\ . \tag{1}$$

It is, however, somewhat of an abstraction, since the actual radiative information resulting from the measurement process is more closely related to an area- and time-average over a finite duration of the order of the time the rain takes to fall to the ground. By contrast, a rain gauge provides a measurement of the time average at a point $\mathbf{x}$ (to the extent a rain gauge can be treated as a geometrical point), over a finite time T:

$$\mathcal{R}_T(\mathbf{x},t) = \frac{1}{T}\int_{-T/2}^{T/2} d\tau\, R(\mathbf{x},t+\tau)\ . \tag{2}$$

It is worth noting that both techniques of measurement are affected by systematic as well as random errors. While remote sensing estimates generally suffer from retrieval errors due to uncertainties in the rain retrieval algorithms used to convert the radiances into rain rates, the rain gauges provide a direct measure of the amount of rain at a location. However, rain gauge data can contain gaps and may be unreliable due to difficulties in continuous monitoring and occasional instrument failure. In addition, rain gauge data are rather sensitive to local factors, such as wind and evaporation, orographic details and placement of the gauges.

Rainfall, despite its apparently irregular nature, is correlated both in space and in time. Spatio-temporal variability of rain directly influences other hydrological processes, such as soil moisture transport and surface run-off. Statistical behavior of area- and/or time-averaged rain rate at different length and time scales can also be an important diagnostic for comparing actual rain behavior with predictions from global climate models (which in general only predict area-averaged rain rate over relatively large grid boxes, typically anywhere from 10 km to several degrees). They also have important consequences for the estimation of the so-called sampling error that arises in satellite measurements because of the intermittent and often incomplete coverage of a given area of the globe [*McConnell and North*, 1987; *Bell et al.*, 1990; *Bell and Kundu*, 1996, hereafter BK96].

In this paper we present a statistical model of space time variability of tropical rainfall seen in precipitation data sets created through ground based radar observations. We describe the data in terms of a simple model used previously by BK96 in their study of the sampling error in preparation for the Tropical Rainfall Measurement Mission (TRMM). In this model Fourier modes of the local instantaneous rain rate $R(\mathbf{x},t)$ are treated as a set of random variables whose evolution is governed by first order linear stochastic equations driven by white noise sources. This equation immediately leads to an analytical form of the rainfall power spectrum depending on a few adjustable parameters. The Fourier transform of the spectrum yields the space-time covariance function of the local rain rate field,

$$c(\mathbf{x},t;\mathbf{x}',t') \;=\; \langle R'(\mathbf{x},t)R'(\mathbf{x}',t')\rangle\ , \tag{3}$$

where

$$R'(\mathbf{x},t) = R(\mathbf{x},t) - \langle R(\mathbf{x},t)\rangle \tag{4}$$

and the angle brackets represent a statistical average over a hypothetical collection, or *ensemble* of rain fields with similar rain climatology. This function when integrated over finite grid box areas or over finite time intervals with respect to each of its space or time arguments, yields the space-time covariance function of the area- or time-averaged rain rate, which are then fitted to the data to determine the model parameters. The model can easily address questions with regard to dependence of the statistical properties of rain on spatial and temporal resolution scales. This model, which in BK96 was fairly successful in describing the Global Atmospheric Research Program (GARP) Atlantic Tropical Experiment (GATE) Phase I data, is here tested with a gridded precipitation data set constructed from

the radar images obtained during the Tropical Ocean Global Atmosphere (TOGA) Coupled Ocean Atmosphere Response Experiment (COARE) [*Webster and Lukas*, 1992]. This data set covers a longer period of observation than the GATE data set studied in BK96 and is gathered from a region in the tropical western Pacific rather than the eastern Atlantic where GATE was conducted. Since the data set gives us instantaneous spatial averages on a regular rectangular cartesian grid analogous to the GATE data, it is suitable for studying the statistical behavior of the rain field averaged over various spatial scales.

There exists a large class of physical rainfall models based on random processes that may be called *hierarchical cluster models.* These models attempt to capture the spatial and temporal organization of storm systems that cause precipitation. A general mathematical framework for such a model was laid by *Le Cam* [1961] in terms of "marked point processes". He used stochastic Poisson processes with suitably randomized parameters to describe the irregular occurrence of rain events in space and time. In many models spatial aggregation is described in terms of an organized hierarchy of structures – "rainbands" of elongated elliptical shape with circular rain "cells" clustered within them. The "marks" refer to a set of physical attributes which may include maximum rainfall intensity, size and speed of the cells and so on. A number of such models of various degrees of complexity exist in the literature [e.g., *Gupta and Waymire*, 1979; *Cox and Isham*, 1988; *Smith and Krajewski*, 1987]. A particularly elaborate stochastic space-time model of precipitation with exponentially decaying gaussian distributed rain intensity field within a rain cell was introduced by *Waymire et al.* [1984] following earlier work by *Gupta and Waymire* [1979]. The Waymire–Gupta–Rodríguez-Iturbe (WGR) model [see *Waymire et al.*, 1984] aims to describe the mesoscale structure of precipitation in some detail, including for instance the effect of global advection of rain bands, at the price of introducing a large number of phenomenological parameters. Estimation of the parameters characterizing a cluster model from actual rain data poses a difficult problem. Parameter estimation for simpler cluster models has been investigated by *Smith and Karr* [1987], *Valdés et al.* [1985], *Smith and Krajewski* [1987] and *Krajewski and Smith* [1989]. *Islam et al.* [1988], *Valdés et al.* [1990], and *Koepsell and Valdés* [1991] used nonlinear least squares fitting to estimate the WGR model parameters – a difficult task in view of the complexities of the model. An analytical form of the model spectrum was derived by *Valdés et al.* [1990]. However, *Valdés et al.* [1994] have found that the WGR model spectrum using the estimated model parameters does not describe the GATE spectrum well. For a recent account of a number of models in this category, including the problem of parameter estimation in the context of a precipitation data set obtained in the United Kingdom from the HYREX project, a large network of radar and rain gauges, see *Wheater et al.* [2000]. It is the ease of parameter estimation and simplicity of application to real rain data sets, combined with its ability to capture the variability on many different space and time scales, that motivated us to consider a rainfall model based on a simple stochastic dynamical equation rather than a cluster model of the type described above.

Other rainfall models based on a stochastic dynamical equation exist in the literature [*North and Nakamoto*, 1989; *Yoo et al.*, 1996] which also describe the GATE spectrum with varying degrees of success. Of these the North-Nakamoto (NN) forced diffusion model turns out to be a special case of the model considered here. The dynamical equation is particularly suited to capturing some essential features such as the diffusive spreading of the rain field through a fluid atmosphere. The underlying stochastic differential equation of the Yoo-Valdés-North (YVN) model includes a second time derivative term which may help in modeling the high frequency behavior of the rainfall spectrum. However, the model has more parameters and has so far been applied only in the low wave number regime, and it is not yet clear how well the model describes rainfall statistics at the smaller spatial scales considered in this paper.

The present model is an outgrowth of previous work by *Bell* [1987] and *Bell et al.* [1990] who constructed a model to simulate the statistics of the gridded rainfall data from GATE. They introduced a gaussian random field on a discrete 4 km grid whose Fourier amplitudes were assumed to obey a stochastic dynamical equation similar to the one used here. A nonlinear mapping was used to generate the gridded rain rate field with a rain-rate probability distribution matching the observed lognormal distribution. Unlike that work, the present model directly describes the stochastic evolution of the instantaneous point rain rate field which can be averaged to any desired length and time scale. Moreover, it yields the full space-time covariance function in terms of a small number of model parameters. Once they are determined from data, the model allows us to relate statistics at different scales, whereas the original model of *Bell* [1987] was applicable only to a single spatial grid scale and, moreover, the form of the spatial covariance function had to be specified as an explicit

function of spatial separation, in effect increasing the number of empirical parameters. However, there is a price to pay – the present model is restricted to discussing only the second moment statistics.

As we shall see in later sections, our model spectrum naturally leads to power law scaling behavior of the rain statistics at small spatial scales suggesting an underlying fractal structure of the rain field. Multifractal rainfall models based on a somewhat different framework have been proposed by *Schertzer and Lovejoy* [1987] and *Gupta and Waymire* [1990] extending earlier ideas of *Lovejoy* [1982], *Lovejoy and Mandelbrot* [1985] and *Waymire* [1985]. These models describe the moments of a spatially averaged rain field in terms of multifractals, which treats the rain field fluctuations (i.e., deviation from the mean) as being governed by an underlying self-similar multiplicative random cascade process.

The paper is organized as follows. Section 2 describes the rainfall model. Here we collect a number of mathematical formulas related to the statistics of area-averaged rain rate field. Of particular interest to us is the way these statistics depend on the averaging length scale. In section 3 we describe the gridded TOGA COARE precipitation data set and the procedure for fitting the model to data and estimating the model parameters. Section 4 is devoted to a discussion of the results focusing, in particular, on the power law dependence of several rain statistics on the averaging length scale perhaps associated with fractal nature of the rain rate field. Some concluding remarks and directions for future work are presented in the final section. Three appendices containing some mathematical details are included.

2. Modeling Rainfall Variability

In this section we describe the framework used in this paper to study rainfall variability. One of our main interests is to investigate how rain statistics depend on the averaging length and time scales. In this paper we concentrate on the statistical properties of rain rate fields spatially averaged over various length scales. Gridded remote sensing precipitation data from ground based or satellite measurements are particularly convenient for such studies, since the gridding allows us to rescale the grid to any desired integral multiple of the original grid size at which the data is supplied, by simply grouping the grid boxes together into larger units. A number of basic statistics are calculated from the data for a range of grid sizes L and compared with model predictions. One of the main ingredients of the model is an analytical form of the rainfall spectrum which is described next.

2.1. A Spectral Model of Rain Rate Covariance

In this subsection we give a brief account of the model. The underlying assumptions of the model as well as many of its salient features have already been described in detail in BK96. A set of results that are useful in fitting the model to data and checking some of its predictions are derived here.

As was noted in *Bell* [1987] and *Bell et al.* [1990], rain rate fields have the characteristic feature that the time scale over which spatial averages remain correlated depends on the size of the averaging area: the larger the averaging area, the longer the time over which the average rain rate remains correlated. This behavior, typical of turbulent fluids, must be captured by a realistic model of rainfall statistics. Although local rain statistics are expected to be affected by inhomogeneous and nonstationary directional effects, such as occur in rain bands, we are interested in a model description that is valid over a large enough region and sufficiently long time for such directional effects to average out, so that the rain statistics can be treated to a good approximation as spatially homogeneous and isotropic and stationary in time. These assumptions imply that the space-time covariance function of the rain field at the points $\mathbf{x}$, $\mathbf{x}'$ and times t,t' has the form

$$c(\mathbf{x},t;\mathbf{x}',t') = c(\boldsymbol{\rho},\tau) = c(\rho,\tau)$$

where $\tau = t'-t$, $\boldsymbol{\rho} = \mathbf{x}'-\mathbf{x}$ and $\rho = |\boldsymbol{\rho}|$. It can be expressed as a Fourier integral

$$c(\boldsymbol{\rho},\tau) = \frac{1}{(2\pi)^{3/2}}\int_{-\infty}^{\infty} d\omega \int d^2\mathbf{k}\, e^{i(\mathbf{k}\cdot\boldsymbol{\rho}-\omega\tau)}\,\tilde{c}(\mathbf{k},\omega)\ , \quad (5)$$

where $\tilde{c}(\mathbf{k},\omega)$ represents the (unnormalized) rain rate power spectrum. Isotropy implies $\tilde{c}(\mathbf{k},\omega) = \tilde{c}(k,\omega)$ where $k = |\mathbf{k}|$. The spatial Fourier amplitudes

$$a(\mathbf{k},t) = \frac{1}{2\pi}\int d^2\mathbf{x}\, e^{-i\mathbf{k}\cdot\mathbf{x}} R'(\mathbf{x},t)$$

of the field $R'(\mathbf{x},t)$ defined by Eq. (4) are assumed to obey the first order linear stochastic equation[1]

$$\frac{d}{dt}a(\mathbf{k},t) = -\frac{1}{\tau_k}a(\mathbf{k},t) + f(\mathbf{k},t)\ . \quad (6)$$

[1]A more general possibility is an equation of the form

$$\frac{d}{dt}a(\mathbf{k},t) = -\int d^2\mathbf{k}'\,\frac{1}{\tau_{kk'}}a(\mathbf{k}',t) + f(\mathbf{k},t)\ ,$$

which allows coupling among different Fourier modes.

Eq. (6) represents what is sometimes referred to as a first order autoregressive [AR(1)] process [*Jenkins and Watts*, 1968]. It is formally analogous to the familiar Langevin equation for the velocity of a particle executing thermal Brownian motion in a viscous fluid. Here $f(\mathbf{k},t)$ is a white noise force satisfying

$$\langle f(\mathbf{k},t)f^*(\mathbf{k}',t')\rangle = (2\pi)F_0\,\delta(\mathbf{k}'-\mathbf{k})\,\delta(t'-t)\,, \quad (7)$$

where in the right hand side we have introduced the Dirac δ-finction, and the asterisk denotes complex conjugation. In Eq. (6) τ_k represents the characteristic relaxation time of $a(\mathbf{k},t)$ taken, as in BK96, to be of the form

$$\tau_k = \frac{\tau_0}{(1+k^2L_0^2)^{1+\nu}}\,, \quad (8)$$

where τ_0 and L_0 are characteristic time and length scales. The length scale L_0 effectively separates the long and the short wavelength regimes of the spatial fluctuations of the rain rate field. As a consequence of Eqs. (6) and (7) the power spectrum has the typical "red noise" form

$$\tilde{c}(k,\omega) = \frac{F_0}{\omega^2+\tau_k^{-2}}\,. \quad (9)$$

The analytic form of τ_k in (8) is suggested by the form chosen by *Bell* [1987] and *Bell et al.* [1990] who, in turn, were motivated by the empirical results of *Laughlin* [1981] from GATE data. As already noted, it contains the North-Nakamoto model as a special case with $\nu = 0$. It may be pointed out that in the original work of *Bell* [1987] and *Bell et al.* [1990], the form of the spatial covariance function was separately introduced from an empirical fit, whereas in the present model it is implicitly contained in the analytic form of τ_k. Because of the fractional exponent, the stochastic dynamical equation governing the evolution of the rain rate field now takes the form of an integro-differential equation rather than a pure differential equation as in the case of NN and YVN models. Physically, it can be thought of as representing a nonlocal "fractal" diffusion process by which rain moves through a turbulent atmosphere. In particular, the negative values of ν, required to fit the data sets studied here appear to better describe the high spectral variability of rain rates at small spatial scales – an important factor in extrapolating statistics to rain gauge scales.

The four model parameters F_0, τ_0, L_0 and ν characterize the second moment statistics. No specific assumption about the actual probability distribution for the random variable $R(\mathbf{x},t)$ is introduced at this stage of the model applications. It should be noted however that the model remains incomplete without such an assumption since the higher order moments cannot be represented within the model. Observationally, area averaged rain rates obey an approximately lognormal or gamma distribution, at least on smaller space and time scales.

Substituting the explicit form of the model spectrum in (5) and performing the integral over ω and the angular integral in the $\mathbf{k}$-plane yields

$$c(\boldsymbol{\rho},\tau) = \int_0^\infty dk\,kJ_0(k\rho)c(k,\tau)\,, \quad (10)$$

where $J_0(x)$ is the Bessel function of the first kind of order zero, and

$$c(k,\tau) = \sqrt{\pi/2}\,F_0\tau_k e^{-|\tau|/\tau_k} \quad (11)$$

represents the lagged covariance of $a(\mathbf{k},t)$ defined by the equation

$$\langle a(\mathbf{k},t)a^*(\mathbf{k}',t')\rangle = (2\pi)c(k,\tau)\,\delta(\mathbf{k}'-\mathbf{k})\,, \quad (12)$$

and is the inverse temporal Fourier transform of $\hat{c}(k,\omega)$. At zero lag ($\tau = 0$) carrying out the remaining k-integration in (10), the spatial covariance of the point rain rate can be expressed in the simple closed form

$$c(\boldsymbol{\rho},\tau=0) = \gamma_0 C_\nu(\rho/L_0)\,, \quad (13)$$

where, for convenience, γ_0 is defined by the relation

$$F_0 = \sqrt{2/\pi}\,\Gamma(1+\nu)(L_0^2/\tau_0)\gamma_0\,, \quad (14)$$

$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$ is the Euler Γ-function, and

$$C_\nu(z) = (z/2)^\nu K_\nu(z) \quad (15)$$

is a function related to the modified Bessel function $K_\nu(z)$. For $z \ll 1$, $C_\nu(z)$ has the behavior

$$C_\nu(z) = \frac{1}{2}\Gamma(\nu) + \frac{1}{2}\Gamma(-\nu)\left(\frac{z}{2}\right)^{2\nu} + O(z^{2+2\nu}) + O(z^2)\,. \quad (16)$$

When $\nu > 0$, the first term dominates and $C_\nu(z) \to \frac{1}{2}\Gamma(\nu)$ as $z \to 0$. In this case the variance of the point rain rate $c(0,0) = \frac{1}{2}\gamma_0\Gamma(\nu)$ is finite. However, as we shall see, like the GATE data, the TOGA-COARE data requires $\nu < 0$ and consequently the covariance function exhibits a power law singularity

$$c(\boldsymbol{\rho},0) \sim \frac{1}{2}\gamma_0\Gamma(|\nu|)(\rho/2L_0)^{-2|\nu|} \quad (17)$$

as the separation $\rho \to 0$. The origin of this singular behavior can be traced to the k-dependence of the characteristic time scale τ_k governing the decay rate of the Fourier mode $a(\mathbf{k},t)$; for the long wavelength modes ($k \approx 0$) τ_k approaches a constant value τ_0, whereas for the short wavelength modes ($k \gg 1/L_0$) τ_k behaves like $k^{-2(1-|\nu|)}$. Note that for values of the exponent ν in the range $-1 < \nu < 0$ (as seems to be the case for both GATE and TOGA-COARE data sets), $\tau_k \to 0$ as $k \to \infty$. For the NN model ($\nu = 0$) the singularity in $c(\boldsymbol{\rho}, 0)$ weakens to become logarithmic.

Next, we turn to the space-time covariance function of area-averaged rain rate,

$$C_{\mathcal{A}\mathcal{A}'}(\mathbf{s},\tau) = \langle \mathcal{R}'_{\mathcal{A}}(t)\mathcal{R}'_{\mathcal{A}'}(t+\tau)\rangle$$

where $\mathcal{A}$, $\mathcal{A}'$ are two $L \times L$ boxes whose centers are separated by $\mathbf{s}$ and the prime on the rain rate variables denotes deviation from the mean, as in Eq. (4). Starting from the formula

$$C_{\mathcal{A}\mathcal{A}'}(\mathbf{s},\tau) = \frac{1}{\mathcal{A}^2}\int_{\mathcal{A}} d^2\mathbf{x}\int_{\mathcal{A}'} d^2\mathbf{x}'\, c(\mathbf{s}+\mathbf{x}'-\mathbf{x},\tau) \quad (18)$$

our spectral model yields at zero lag ($\tau = 0$) the expression

$$\begin{aligned} C_{\mathcal{A}\mathcal{A}'}(\mathbf{s},0) &= 4\gamma_0 \int_0^1\int_0^1 d\xi_1\, d\xi_2\, (1-\xi_1)(1-\xi_2) \\ &\times\ C_\nu\left(\left|\boldsymbol{\xi} + \frac{\mathbf{s}}{L}\right|\frac{L}{L_0}\right) . \end{aligned} \quad (19)$$

See BK96 Appendix A for details of the derivation. Upon setting $\mathbf{s} = 0$ in Eq. (18), the variance of the area-averaged rain rate in a grid box $\mathcal{A}$ of area $A = L^2$ is given by

$$\begin{aligned} \sigma_A^2 &\equiv \sigma^2(L) = C_{\mathcal{A}\mathcal{A}'}(\mathbf{s}=0,\tau=0) \\ &= 4\gamma_0 G(\nu; L/L_0) , \end{aligned} \quad (20)$$

where $G(\nu; z)$ is the double integral

$$\begin{aligned} G(\nu;z) &= \int_0^1\int_0^1 d\xi_1\, d\xi_2\, (1-\xi_1)(1-\xi_2) \\ &\times\ C_\nu\left(z\sqrt{\xi_1^2+\xi_2^2}\right) . \end{aligned} \quad (21)$$

The power law singularity of the spatial covariance of the point rain rate as a function of separation in turn leads to a simple scaling behavior of the statistics of area averaged rain with the size of the area. Using the first two terms of the expansion (16) in Eq. (21), which dominate as long as $-1 < \nu < 0$, we have

$$G(\nu;z) \approx \frac{1}{8}\Gamma(-|\nu|) + \frac{1}{2}\Gamma(|\nu|)H(\nu)z^{-2|\nu|} , \quad (22)$$

where $H(\nu)$ denotes the integral

$$H(\nu) = \int_0^1\int_0^1 d\xi_1\, d\xi_2\, (1-\xi_1)(1-\xi_2)(\xi_1^2+\xi_2^2)^\nu . \quad (23)$$

Hence, for small L (more precisely $L \ll L_0$), $\sigma^2(L)$ can be expressed in the form

$$\sigma^2(L) \approx a_0 + b_0 L^{-2|\nu|} , \quad (24)$$

where a_0 and b_0 are constants given by

$$a_0 = \frac{1}{2}\gamma_0\Gamma(-|\nu|) , \quad (25)$$

$$b_0 = 2\gamma_0\Gamma(|\nu|)H(\nu)(2L_0)^{2|\nu|} . \quad (26)$$

Thus the model predicts that as $L \to 0$, the variance of area-averaged rain rate $\sigma^2(L)$ exhibits a power law singularity:

$$\sigma^2(L) \sim L^{-2|\nu|} .$$

We note that although the point variance is infinite in the model, it is not necessarily inconsistent with experience since observations always involve smoothing of the point rain rate field in space and/or time, and variances of smoothed rain field are finite in the model.

The lagged correlation function

$$\Phi_{\mathcal{A}\mathcal{A}'}(\mathbf{s},\tau) = \frac{C_{\mathcal{A}\mathcal{A}'}(\mathbf{s},\tau)}{\sigma_A^2} \quad (27)$$

is a rapidly decreasing function of the separation $\mathbf{s}$. The model correlation function is markedly nonexponential. Its fall-off rate is conveniently characterized by the "integral correlation length" $\Lambda_{\rm int}(L)$ [which would be the $(1/e)$-folding distance for an exponentially decaying correlation function] defined in the isotropic case through the relation

$$\Lambda_{\rm int}^2(L) = \int_0^\infty ds\, s\Phi_{\mathcal{A}\mathcal{A}'}(\mathbf{s},\tau=0) . \quad (28)$$

It is related to the model spectrum through

$$\Lambda_{\rm int}^2(L) = \frac{1}{\sigma^2(L)}c(\mathbf{k}=0,\tau=0) , \quad (29)$$

which upon using Eq.(20) reduces to

$$\Lambda_{\rm int}^2(L) = \frac{1}{4}\Gamma(1+\nu)L_0^2/G(\nu; L/L_0) . \quad (30)$$

From Eq. (22) it then follows that $\Lambda_{\rm int}(L)$ vanishes as $L^{|\nu|}$ as $L \to 0$.

We can also define an integral autocorrelation time $\tau_{\rm int}(L)$ (which would be the $1/e$-folding time for an exponentially decreasing autocorrelation function) through the formula

$$\tau_{\rm int}(L) = \int_0^\infty d\tau\, \Phi_{\mathcal{A}\mathcal{A}}(\mathbf{s}=0,\tau)\,. \tag{31}$$

It can be reexpressed in the form

$$\tau_{\rm int}(L) = \sqrt{\frac{\pi}{2}}\frac{\tilde{C_{\mathcal{A}}}(\omega=0)}{\sigma_A^2}\,, \tag{32}$$

where $\tilde{C_{\mathcal{A}}}(\omega)$ denotes the Fourier transform of the lagged autocovariance function $C_{\mathcal{A}\mathcal{A}}(\mathbf{s}=0,\tau)$. Evaluation of the integral in the numerator (see Appendix A) yields

$$\tau_{\rm int}(L) = \tau_0 \frac{\Gamma(1+\nu)}{\Gamma(2+2\nu)}\frac{G(1+2\nu;L/L_0)}{G(\nu;L/L_0)}\,. \tag{33}$$

The values of ν that best describe rain data in fact appear to lie in the narrower range $-1/2 < \nu < 0$ (rather than the range $-1 < \nu < 0$ considered above). The numerator of Eq. (33) is then finite and consequently the model predicts that $\tau_{\rm int}(L) \to 0$ as $L^{2|\nu|}$ as $L \to 0$. In other words, the "time scale" for fluctuations of average rain rate in a box area L^2 gets smaller with decreasing box size.

The case $\nu = -1$ presents a particularly simple and analytically tractable special case of our spectral model. In this case, the characteristic length L_0 drops out of the model. All the spatial Fourier modes excited by the random white noise source decay exponentially with the same characteristic e-folding time τ_0. The Fourier transform of the spatial covariance $c(\boldsymbol{\rho},0)$, namely, $c(\mathbf{k},0)$ is a constant, which in turn implies that the point rain rate field at equal times is δ-correlated:

$$c(\boldsymbol{\rho},0) \propto \delta(\boldsymbol{\rho}).$$

In this case the covariance function of the area-averaged rain rate at zero lag can be easily computed for $\nu = -1$ to yield the explicit formula

$$\begin{aligned} C_{\mathcal{A}\mathcal{A}'}(\mathbf{s},0) &= \left(\frac{\text{const}}{L^2}\right)\left(1-\frac{|s_1|}{L}\right)\left(1-\frac{|s_2|}{L}\right) \\ &\times\ W\left(\frac{|s_1|}{L}\right)W\left(\frac{|s_2|}{L}\right)\,, \end{aligned} \tag{34}$$

where $W(x)$ is the window function

$$W(x) = \begin{cases} 1 & |x|<1 \\ 0 & |x|>1 \end{cases}\,.$$

The correlation between the rain rates in disjoint grid boxes vanish: the integral correlation length $\Lambda^2_{\rm int}(L)$ is simply equal to the box size L. The variance of area-averaged rain rate $\sigma^2(L)$ is proportional to $1/L^2$.

2.2. Some Other Rain Statistics

The spectral model presented above makes a number of predictions about the change of rain statistics with scale. With a few additional assumptions these predictions can be shown to have interesting consequences for the asymptotic behavior of the "conditional" statistics of rain as well. We now turn to a brief description of these statistics.

Consider remote sensing measurements of precipitation over a given area and time period and the resulting rain maps gridded on an $L \times L$ square grid. The ensemble mean rain rate $\langle R\rangle$ is estimated as the average of all the data within the data set, and, due to the assumption of statistical homogeneity, is independent of the resolution scale L. However, other statistical properties of the field, such as the fraction of grid boxes containing nonzero rain ("the rainy area fraction") $p(L)$ and the variance of the grid box-averaged rain rate $\sigma^2(L)$ are expected to depend nontrivially on L. Two other related statistics are the mean $r_c(L)$ and the variance $\sigma_c^2(L)$ of the grid box-averaged rain rate conditional on the box containing non zero rain. It is easy to show that (see, e.g., BKK, Appendix A) these are related through

$$\langle R\rangle = p(L) r_c(L), \tag{35}$$

$$\sigma^2(L) = p(L)[\sigma_c^2(L) + r_c(L)^2] - \langle R\rangle^2 \tag{36}$$

It is convenient to define the dimensionless ratio

$$\mu \equiv \mu(L) = \sigma_c(L)/r_c(L). \tag{37}$$

Using Eq. (35), Eq. (36) can then be written as

$$1 + \frac{\sigma^2(L)}{\langle R\rangle^2} = \frac{1+\mu^2}{p(L)}. \tag{38}$$

Based on analysis of rain gauge data from Darwin, Australia and Melbourne, Florida and radar data from GATE, *Short et al.* [1993] have suggested that μ is relatively constant over a range of length scales, averaging times, types of data (radar or rain gauge) and rain climates. In particular, μ is insensitive to the averaging length scale L. This was also corroborated by the results of BKK from precipitation data obtained from SSM/I instruments on board NOAA DMSP satellites F10 and F11.

If the quantity $\mu(L)$ remains finite as $L \to 0$, then Eq. (38) together with the asymptotic behavior (24) predicts that $p(L)$ would vanish as $L^{2|\nu|}$ in this limit. The limiting behavior $p(L) \to 0$ as $L \to 0$ in turn implies that the local instantaneous rain rate field $R(\mathbf{x},t)$ vanishes *almost everywhere*, i.e., everywhere except a

set of measure zero – a familiar property of spatial fractals. We shall return to this in greater detail in section 4 when we present the results of our data analysis.

3. Testing the Model Predictions with Ground based Radar Data

3.1. The TOGA COARE dataset

The radar data we use to test the spectral model described above were collected during TOGA COARE [*Webster and Lukas*, 1992] carried out in the Western Pacific Equatorial Warm Pool during the period November 1992 to February 1993. The measurements were made with a pair of calibrated Doppler radars (labeled TOGA and MIT) on board two research ships located near the center of the Intensive Flux Array (IFA) – a region in the shape of a quadrilateral, roughly centered at 2° S, 156° E. Radar returns weaken with distance from the radar, and quantitatively reliable rain rate estimates were provided out to a distance 145 km from each radar. The fields of view (FOVs) of the two radars overlapped substantially when they were both operating. The measured radar reflectances were converted into average surface rain rates on a 2 km spatial grid using an empirical $Z - R$ relation. The gridded rainfall data set used in the present work is described by *Short et al.* [1997]. The total 101-day period of observation was divided into three "monthly" cruises covering the periods November 11 to December 10 1992, December 15 1992 to January 18 1993 and January 23 to February 23 1993 designated respectively as Cruises 1, 2 and 3. Rain maps were available roughly at intervals of ten minutes with occasional long gaps during which data from at least one of the radars were missing.

3.2. Statistical Analysis of the Data

The gridded data set obtained from the TOGA COARE observations yields estimates of the various area-averaged rain statistics introduced previously, allowing one to test the model predictions. We now briefly describe how these statistics are obtained from the data.

The statistics are derived using rain estimates from square areas 128×128 km^2 centered at the location of each of the two ships. The locations of the areas changed slowly with time because of the slight drift of the ships during the month-long experiments. However, the effect of this on the statistics, if any, is expected to be negligible. Each area is subdivided into boxes $L \times L$ in size, with $L = 2^n$ km ($n = 1, 2, \ldots, 7$), and the quantities $\langle R \rangle$, $\sigma^2(L)$, $p(L)$, $r_c(L)$ and $\sigma_c(L)$ defined in the previous section are computed for each L from the data. Only those $L \times L$ boxes in a rain map were used for which at least 95% of the box had valid data. This was necessary since the radar images frequently had data missing along a radial line emanating from the center of the circular FOV where the radar is located. This data dropout phenomenon affected the grid boxes located close to the origin through which this radial line passed, especially at the smaller grid length scales $L = 2, 4, 8$ and 16 km.

For comparison purposes, we also computed the same spatial statistics for the GATE Phase I (1716 images) and Phase II (1512 images) data from the 280 km box centered at the location of the ship. For GATE data the basic grid size is 4 km.

We determined the integral correlation time $\tau_{\rm int}(L)$ for each L as follows. Mean rain rate time series were constructed for each $L \times L$ box within the selected areas for the three cruises. Again only boxes with at least 95% of the area containing data were included in the time series in order to contend with the data loss phenomenon mentioned above. The lagged autocovariance function $C_{\mathcal{A}\mathcal{A}}(0, \tau)$ averaged over all boxes in the chosen areas was computed for each L at lags τ of multiples of 10 min.

As an example, the lagged autocorrelation function

$$\Phi_L(\tau) \equiv \Phi_{\mathcal{A}\mathcal{A}}(0, \tau)$$

for $L = 128$ km for the six data sets are shown in Figure 1. For each data set a numerical estimate of the time integral $\tau_{\rm int}(L) = \int_0^\infty \Phi_L(\tau) d\tau$ was obtained. This is discussed in more detail later.

The following subsection describes how the parameters γ_0, L_0, ν and τ_0 are determined by fitting the data to the statistics of gridded averages predicted by the model.

3.3. Estimation of the Model Parameters

The parameters γ_0, L_0 and ν determine the spatial statistics at zero lag. They are estimated by fitting the variance of the gridded rain rate average $\sigma^2(L)$ to the asymptotic formula (24) valid for small L. The fit determines the exponent ν and the coefficients a_0 and b_0. Eq. (25) then determines the parameter γ_0 and Eq. (26) finally determines the characteristic length scale L_0. The procedure was carried out for the TOGA COARE as well as both GATE Phase I and Phase II data sets.

An estimate of the characteristic time τ_0 for each data set can be obtained by averaging over the values individually computed by solving (33) for τ_0 using the

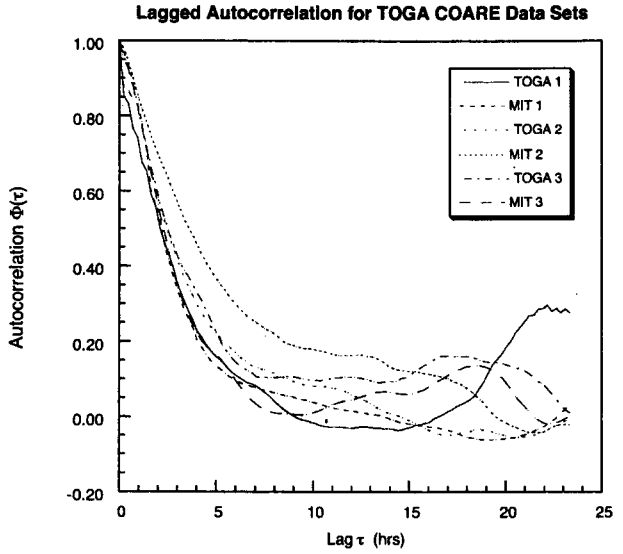

Figure 1. Plot of lagged autocorrelation functions $\Phi_L(\tau)$ for $L = 128$ km square boxes centered at the location of the TOGA and MIT radars for each of the three cruises.

values of the integral correlation time $\tau_{\rm int}(L)$ for various box sizes L and the previously determined parameter values of ν and L_0. Since there is only a finite length T of data available, one cannot actually calculate $\tau_{\rm int}(L)$ from the data itself. The data can only supply an estimate of (31) with a finite cut-off,

$$\hat{\tau}_{\rm int}(L, \tau_0; \tau_{\rm max}) = \int_0^{\tau_{\rm max}} \Phi_L(\tau) d\tau \qquad (39)$$

with $\tau_{\rm max}$ less than T. In principle, we could solve (39) for τ_0 using the model from the data-derived value of $\hat{\tau}_{\rm int}$ for a chosen cut-off $\tau_{\rm max}$. In practice this is difficult, since the model-predicted τ_0-dependence of the function $\hat{\tau}_{\rm int}(L, \tau_0; \tau_{\rm max})$ is nonlinear and complex to calculate. Instead, we adopt an iterative method which we describe next.

There is an uncertainty inherent in choosing $\tau_{\rm max}$ in order to get the "best estimate" of τ_0. It arises because the tail of the autocorrelation function suffers from a combined effect of statistical noise and possible secular variations of the rainfall pattern over longer time scales not represented in our spectral model. The choice of the cut-off is guided by several considerations. On the one hand, $\tau_{\rm max}$ should be large enough so that the "residual" effect of the tail in the integral (31) is small. Some guidance on an acceptable value of $\tau_{\rm max}$ is provided by the model's prediction of the the truncated integral. However, the empirical lagged autocorrelation suffers from sampling error because it is obtained from a finite time series of length $T \approx 30$ days. The cut-off $\tau_{\rm max}$ must therefore satisfy the condition $\tau_{\rm max} \ll T$. Also $\tau_{\rm max}$ should be chosen to be small enough so that the effect of slow secular modulations of the lagged autocorrelation function caused by physical phenomena such as a diurnal cycle or various atmospheric waves, which are not accounted for in our simple model spectrum, is minimized. Evidence of such longer time scales was most obvious in the Cruise 2 MIT and the Cruise 3 TOGA radar data sets. These modulations can probably be attributed to various large scale atmospheric waves known to influence rainfall in the equatorial Pacific. We shall revisit some of these issues in the next section. Keeping these factors in mind, the cut-off parameter $\tau_{\rm max}$ employed in our computations was set at 12 hrs.

The choice of the upper limit of integration is somewhat subjective. The tails of the observed correlation functions were substantial for several of the data sets, most notably, TOGA Cruise 3, for which it extended well beyond 20 hours, at which lags one might expect effects of diurnal variation and atmospheric waves to influence the statistics. We therefore chose a relatively small cut-off $\tau_{\rm max} = 12$ hours with the assumption that the spectral model accurately describes the observed correlations up to the lag $\tau = \tau_{\rm max}$.

Next, we turn to the problem of estimation of τ_0. Using (39), Eq. (31) for the integral correlation time $\tau_{\rm int}(L) \equiv \tau_{\rm int}(L, \tau_0)$ can be expressed as the sum

$$\tau_{\rm int}(L, \tau_0) = \hat{\tau}_{\rm int}(L, \tau_0; \tau_{\rm max}) + \int_{\tau_{\rm max}}^{\infty} \Phi_L(\tau) d\tau \qquad (40)$$

If the left-hand side of (40) were known, Eq. (33) would immediately yield τ_0. However, while the first term on the right-hand side of (40) is obtained from data, the second term can be calculated from the model only when τ_0 is known. We therefore solve Eq. (40) iteratively: to begin the process, we obtain a first guess for τ_0 by ignoring the contribution of the integral; for all the following steps of the iteration we evaluate the integral on the right-hand side using the model $\Phi_L(\tau)$ with the value of τ_0 obtained from the previous step. For ease of computation, we replace the square box by a circular disk of equal area. This approximation is sufficient for our purposes. Appendix B gives some details of the model calculations for a circular area. In practice, for the TOGA COARE data, the correction to τ_0 due to the introduction of a cutoff is of order 5%, and it was unnecessary to go beyond the first iteration.

The iterative method applied to GATE Phase I data with $\tau_{\rm max}$ chosen to be 18 h gives an estimate of τ_0

within about 1% of what was reported in BK96 from a direct fit to the lagged autocorrelation function.

4. Results and Discussion

The full set of parameter values for which the model best fits the TOGA COARE data are given in Table 1, along with the corresponding values for GATE Phase I determined in BK96 using a somewhat less exact (and more arduous) trial-and-error curve fitting method instead of the analytical method adopted here.

4.1. Model Parameters ν, L_0 and γ_0

Table 1 summarizes the TOGA COARE model parameters. We see that the values of the power law exponent ν range between about -0.21 to -0.34 indicating a much stronger singular behavior than was found for the GATE Phase I data ($\nu = -0.11$). The characteristic time scales τ_0 lie in the range of roughly 4.5 to 8.2 hours, considerably shorter than the GATE Phase I value $\tau_0 \approx 13$ hours. The characteristic length L_0 ranges between about 54 and 94 km indicating a somewhat sharper fall-off of the spatial correlations compared to GATE Phase I for which $L_0 = 104$ km. The results exhibiting the quality of the fit are shown in Figure 2 for the TOGA and MIT radars. It is seen that in most cases the fit is quite good up to grid size $L = 64$ km, beyond which the approximation on which Eq. (24) is based is expected to break down. Our results seem to be consistent with the model prediction that the variance $\sigma^2(L)$ has a power law behavior $\sigma^2(L) \sim L^{-2|\nu|}$ as $L \to 0$. Robustness of the parameterization was checked by comparing the variance of area-averaged rain rate $(\sigma_A^2)_{\text{th}}$ for an area $A = L^2$ with $L = 128$ km predicted by the exact formula (20), rather than the asymptotic formula (24), against the observed variance $(\sigma_A^2)_{\text{obs}}$ for the 128 km box centered at the two ship locations. Table 2 shows the comparisons, which are quite good.

For GATE data we find that for L between 4 km and 56 km, the variance $\sigma^2(L)$ is quite accurately fitted (not shown here) by the formulas

$$\sigma^2(L) = \begin{cases} 17.3L^{-0.24} - 4.66 \text{ mm}^2\text{h}^{-2} & \text{Phase I} \\ 15.0L^{-0.38} - 1.92 \text{ mm}^2\text{h}^{-2} & \text{Phase II} \end{cases},$$

which through Eqs. (24) to (26) yield the parameter values

$$\gamma_0 = 1.02 \text{ mm}^2\text{h}^{-2},\ \nu = -0.12,\ L_0 = 93.8 \text{ km}$$

for Phase I and

$$\gamma_0 = 0.63 \text{ mm}^2\text{h}^{-2},\ \nu = -0.19,\ L_0 = 82.7 \text{ km}$$

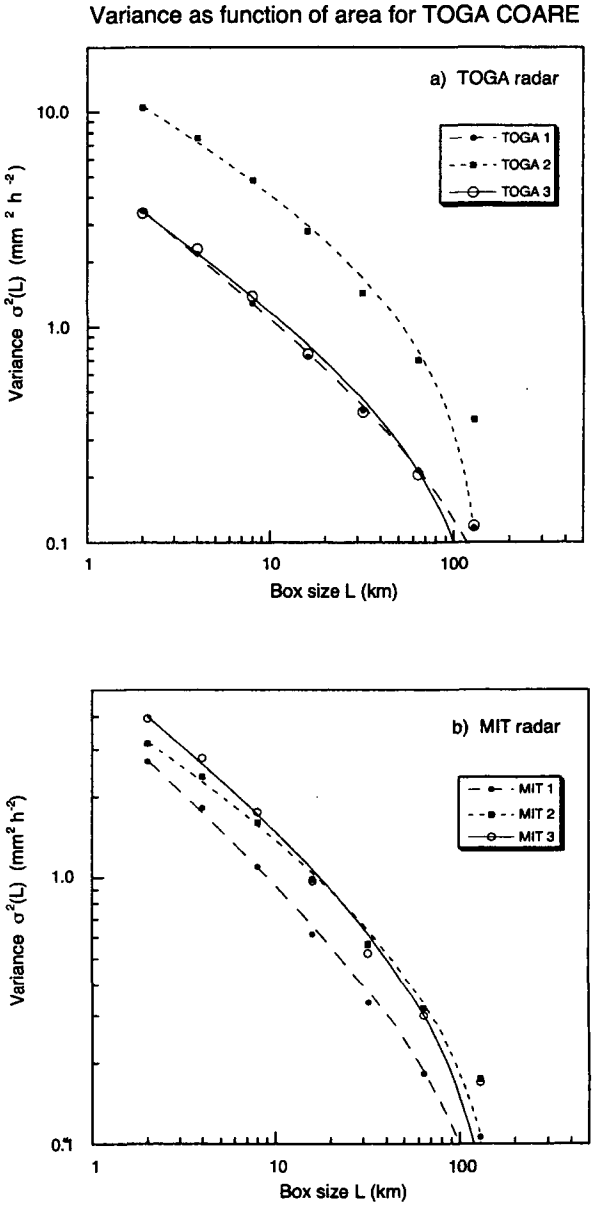

Figure 2. Comparison between the observed variance of area-averaged rain rate $\sigma^2(L)$ for the TOGA COARE data and the corresponding spectral model predictions for various box sizes L ranging from 2 km to 128 km using Eqs. (24)–(26) and the parameter values listed in Table 1.

Table 1. Parameter Values for Rain-Rate Covariance Model.

Dataset	γ_0 ($mm^2\,h^{-2}$)	ν	L_0 (km)	τ_0 (h)	$\langle R\rangle$ ($mm\,h^{-1}$)
GATE Phase I	1.0	−0.11	104.	13.0	0.492
TOGA Cruise 1	0.067	−0.335	94.06	6.8	0.139
MIT Cruise 1	0.086	−0.297	73.89	5.8	0.134
TOGA Cruise 2	0.616	−0.239	53.81	5.0	0.351
MIT Cruise 2	0.206	−0.205	70.40	8.2	0.229
TOGA Cruise 3	0.127	−0.290	61.04	5.2	0.155
MIT Cruise 3	0.180	−0.259	64.94	4.5	0.200

Parameters for the model spectrum obtained from fits to radar data from Phase I of GATE and from the ships MIT and TOGA during the three TOGA COARE cruises. Average rain rate for each dataset is given in the last column.

for Phase II. [The original GATE Phase I parameters listed in Table 1 lead to the asymptotic formula $\sigma^2(L) = 16.80L^{-0.22} - 4.89$ mm^2h^{-2} (see BK96 Appendix A), whose coefficients and the power law exponent are well within the 95% confidence interval of the corresponding quantities in the empirical least squares fit given above.]

We also explored the dependence of the model parameters on the mean rain rate $\langle R\rangle$. While no systematic dependence could be discerned for L_0, τ_0 and ν, the strength parameter γ_0 was found to depend on $\langle R\rangle$ according to the simple formula

$$\gamma_0 = \kappa\langle R\rangle^2, \tag{41}$$

where κ is a dimensionless constant. The fit in Figure 3, which also includes points corresponding to GATE Phase I ($\langle R\rangle = 0.492$ $mm\,h^{-1}$) and GATE Phase II ($\langle R\rangle = 0.386$ $mm\,h^{-1}$), yields a value $\kappa \approx 4.33$.

4.2. Model Parameter τ_0

Next, we turn to a discussion of the dependence of the characteristic time scale $\tau_{\rm int}(L)$ on the spatial averaging scale L. The plots in Fig. 4 show reasonable agreement between the values of $\tau_{\rm int}(L)$ determined from data and the corresponding theoretical predictions computed for various box sizes L using the model parameters. Linearity of the log-log plot for small L confirms the power law dependence $\tau_{\rm int}(L) \sim (L/L_0)^{2|\nu|}$ as $L \to 0$ expected from our model. In BK96 it was found that the integral correlation time estimates for the GATE Phase I is described quite well by the power law form $\tau_{\rm int}(L) = 0.67L^{0.49}$ h. Again the exponent of the power law is rather different from the expected value 0.22 obtained from the spectral model fit. However, it

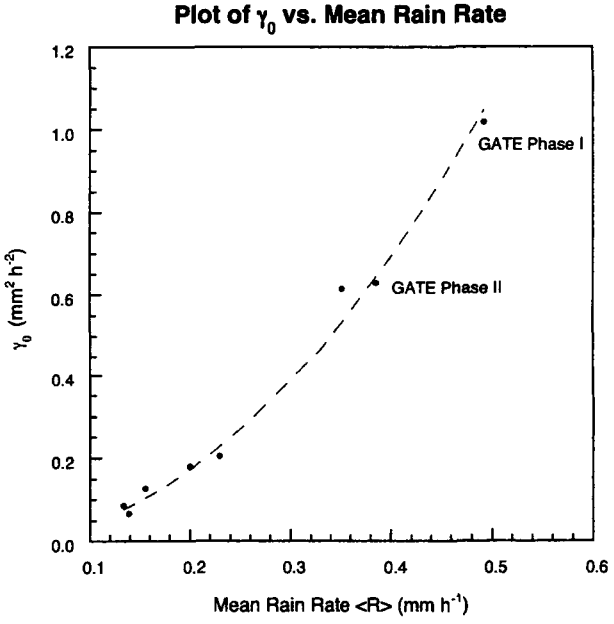

Figure 3. Plot of the parameter γ_0 against the mean rain rate $\langle R\rangle$ for the TOGA COARE and GATE data sets. The continuous curve shows the least squares fit to the formula (43)

Table 2. Comparison between theoretical and observed values of σ_A^2 for $L = 128$ km.

Dataset	$(\sigma_A^2)_{\rm th}$ (mm^2 h^{-2})	$(\sigma_A^2)_{\rm obs}$ (mm^2 h^{-2})
TOGA Cruise 1	0.107	0.116
MIT Cruise 1	0.093	0.105
TOGA Cruise 2	0.399	0.371
MIT Cruise 2	0.176	0.174
TOGA Cruise 3	0.107	0.119
MIT Cruise 3	0.155	0.169

Comparison between values of the variance of area-averaged rain rate σ_A^2 computed from the model using Eq. (20) and the parameters listed in Table 1 for $L = 128$ km and from the data for two 128 km boxes centered at each of the two ship locations.

was also found there that while the model captures the overall decay of the autocorrelation function $\Phi_L(\tau)$, it substantially underestimates the correlation at smaller lags, indicating a departure from the model spectrum for a first order process. Indeed, it was shown by *Bell and Reid* [1993] that this behavior could be accounted for by going over to a second order model, which of course introduces more adjustable parameters.

A number of caveats in the estimates of τ_0 should be pointed out. A fundamental uncertainty in the value of $\tau_{\rm int}(L)$ arises from the fact that it must be estimated from a time series of finite length T which in our case is of the order 30 days. It is known [*Bell*, 1980] that if the true autocorrelation function of a normally distributed random variable has an exponentially decaying form $\Phi_L(\tau) = \exp(-\tau/\tau_{\rm int})$ then the integral correlation $\hat{\tau}_{\rm int}$, defined by Eq. (38) with a finite cut-off $\tau_{\rm max}$, has variance

$$\sigma^2[\hat{\tau}_{\rm int}] = 4(\tau_{\rm max}/T)\hat{\tau}_{\rm int}^2, \qquad (42)$$

where it is assumed that $\tau_{\rm int}(L) \ll \tau_{\rm max} \ll T$. A proof of Eq. (42) is given in Appendix C. Thus, in order to reduce the statistical uncertainty in the estimate, $\tau_{\rm max}$ should be chosen to be as small as possible but large enough to include the entire range over which $\Phi_L(\tau)$ differs appreciably from zero. In our case, if we choose $\tau_{\rm max} \approx 12$ hrs, then it follows that the 2σ uncertainty in the estimate of $\tau_{\rm int}(L)$ obtained from a single time series could, in fact, be as large as $0.5\hat{\tau}_{\rm int}$. It is to be noted that, strictly speaking, this estimate holds only for a gaussian random variable undergoing a linear stochastic dynamical process characterized by an exponentially decaying autocorrelation function. However, the probability distribution of rain rate variable is markedly non-gaussian and the autocorrelation functions possess a much more slowly decaying tail. These factors can lead to an uncertainty larger than what is predicted by Eq. (42). On the other hand, for the smaller grid boxes of size $L = 2^n$ km with $n \leq 6$, the actual uncertainty should be considerably lower than the estimate (42) since the autocovariances were obtained by averaging over the results for the time series for the 4^{7-n} boxes. The time series are, however, not completely independent of one another, because of the spatial correlation of the rain rate in a box with others in its vicinity, and an exact estimate of the uncertainty would be difficult to obtain.

4.3. Conditional Statistics and Spectral Model

The spectral model described in section 2 allows one to quantify how the statistics of area-averaged rain rate depends on the averaging scale L. These statistical quantities, such as the variance $\sigma^2(L)$ and the integral correlation length and time scales, pertain to variability of the rain field including its zeroes, i.e., the region where it is not actually raining. However, unlike most random variables in nature, the instantaneous rain rate has a finite *nonzero* probability of having the sharp value zero, in accordance with what is a matter of common experience – namely that it is not raining in most places most of the time. Mathematically this is expressed by the property that the probability distribution function for the area-averaged instantaneous rain rate $\mathcal{R}_{\mathcal{A}}(t)$ defined by (1) has a δ-function peak at zero followed by an (approximately lognormal) continuum [*Kedem et al.*, 1990]. The probability $p(L)$ for the value $\mathcal{R}_{\mathcal{A}} \neq 0$ is expected to have nontrivial L-dependence

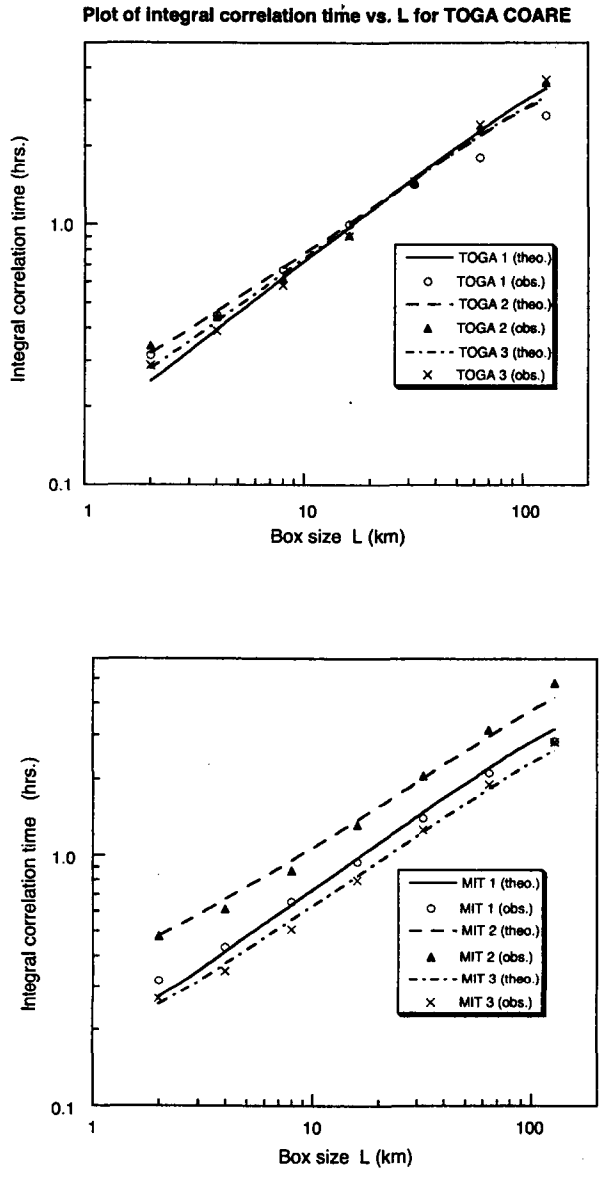

Figure 4. Plot of the integral correlation time $\hat{\tau}_{\rm int}$ evaluated with a finite cut-off $\tau_{\rm max}$ for the three TOGA and MIT cruises and various box sizes L ranging from 2 km to 128 km. The curves represent interpolation of the values computed from the spectral model.

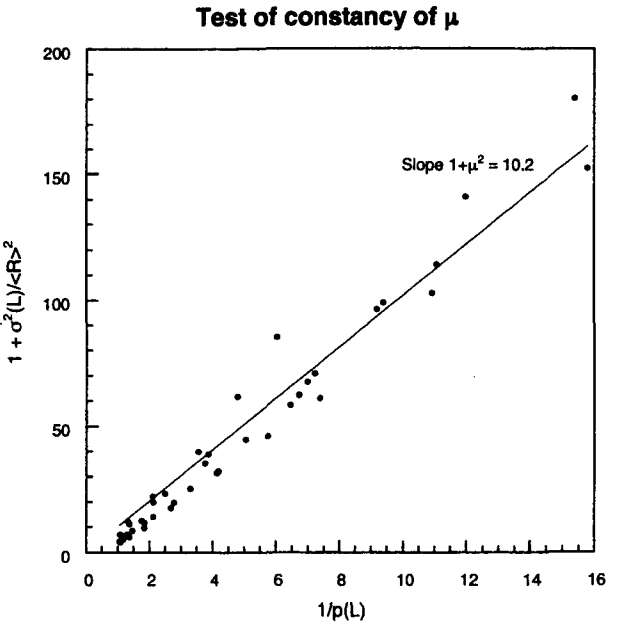

Figure 5. Scatterplot of $1 + \sigma^2(L)/\langle R\rangle^2$ vs. $1/p(L)$ for all the six TOGA COARE data sets for averaging length scales L ranging from 2 km to 128 km ($6\times 7 = 42$ data points). Least squares fit to a straight line through the origin yields the slope $1 + \mu^2 \approx 10.2$, i.e. $\mu \approx 3.03$ for the entire TOGA COARE data set.

and is an important statistic of the gridded rain field. As discussed in section 2, an additional property of the conditional statistics of rain, namely the constancy of the dimensionless ratio μ, allows one to draw some inferences about the asymptotic L-dependence of the function $p(L)$. We now turn to a discussion of this statistic for the TOGA COARE data set. It is, however, worth pointing out that in doing so we are exploring territory that lies outside the scope of the original spectral model.

Our data set provides a simple way to test the constancy of the ratio μ [defined by Eq. (37)] as suggested by *Short et al.* [1993]. Eq. (38) implies that the plot of $1 + \sigma^2(L)/\langle R\rangle^2$ vs. $1/p(L)$ would be a straight line through the origin if μ is a constant independent of L. A scatterplot of all the data points from the six data sets (2 ships, 3 cruises) and each value of L is shown in Figure 5. The linear fit seems satisfactory; the slope yields a mean value of $\mu \approx 3.03$ collectively for all the six data sets. However, there is a pronounced systematic deviation from the linear fit in the "large" $p(L)$ regime [i.e., $p(L) \simeq 1$], which corresponds to large L. One can also evaluate μ separately for each data set.

The individual values were found to be within a narrow range, with a low value of 2.63 for MIT Cruise 2 and a high value of 3.39 for TOGA Cruise 2. As noted in section 2, constancy of μ (or more generally, finiteness of μ as $L \to 0$) together with (24) leads to another important conclusion – a power law scaling behavior of the rain probability $p(L)$ at small L. Combining the formulas (24)–(26) and (41) with Eq. (38) we infer that for $L \ll L_0$,

$$1/p(L) \propto 1 + \beta(L/2L_0)^{-2|\nu|}, \tag{43}$$

where

$$\beta = \frac{2\kappa\Gamma(|\nu|)H(\nu)}{1 + \frac{1}{2}\kappa\Gamma(-|\nu|)}, \tag{44}$$

$H(\nu)$ is given by Eq. (23) and κ is the dimensionless number introduced in (41). This immediately implies the power law behavior

$$p(L) \sim \left(\frac{L}{L_0}\right)^{2|\nu|} \tag{45}$$

as $L \to 0$, which can be independently checked from the data. Fits to a power law $p(L) \propto L^{\eta}$ for each of the six data sets are shown in Figure 6. A straight line fit on a plot of $\log p(L)$ vs. $\log L$ provides evidence of a power law behavior between $L = 2$ km and 64 km. In Table 3 the observed exponent $\eta_{\rm obs}$ is compared against the theoretically expected value $\eta_{\rm th} = 2|\nu|$ for each data set. We see that, although there is generally speaking a close agreement between the two sets of values, for four of the six data sets $\eta_{\rm obs}$ is somewhat larger than $\eta_{\rm th}$ contrary to what one would infer from Eq. (38). However, it should be noted that according to the model the power law dependence on L becomes accurate only in the scaling regime $L \ll L_0$. Consequently, it is more appropriate to compare $\eta_{\rm th}$ with the power law exponent determined from $p(L)$ in the small L limit. Indeed, a closer inspection of the graphs of $p(L)$ shows a slight but systematic deviation at small L from the power law obtained by least squares fit, indicating a smaller slope there. In Table 3 we have also included an estimated exponent $\eta^0_{\rm obs}$ evaluated from the slope of a linear interpolation between $L = 2$ km and 8 km, which clearly improves agreement between the model and observation. One can also see in retrospect that the constancy of μ which gave rise to the power law scaling behavior of $p(L)$, itself becomes progressively more accurate in the limit of small L.

For the GATE data we find that within the range of L between 4 km and 56 km the rain probability $p(L)$ (L in km) is represented quite well by the power law

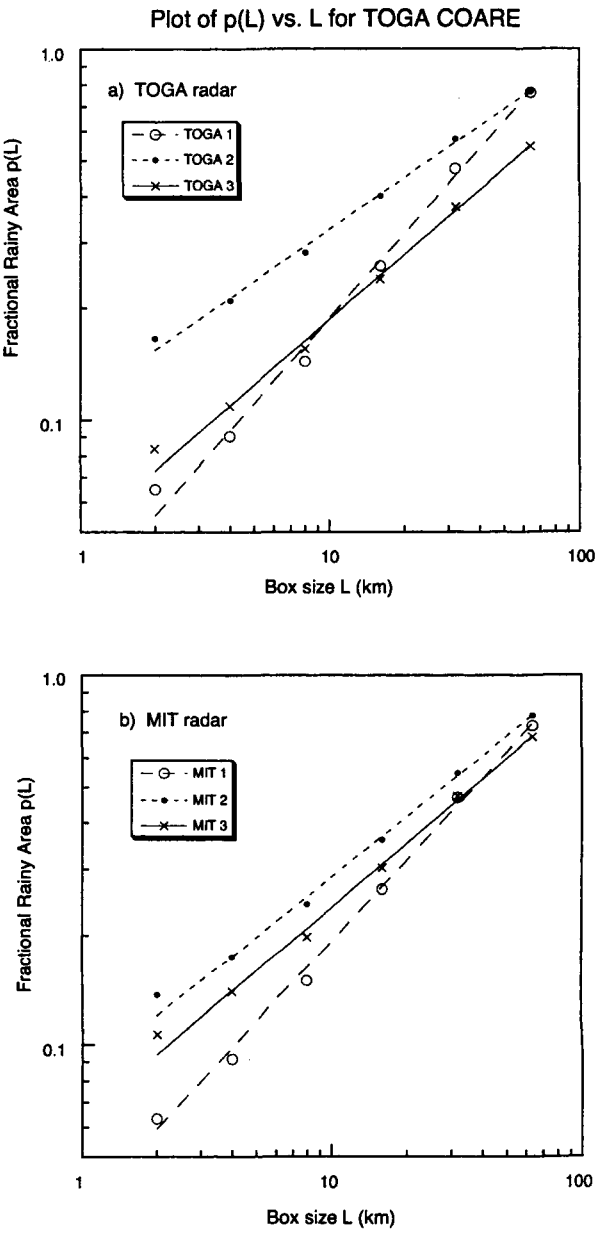

Figure 6. Power law fit to the fractional rainy area $p(L)$ for the three TOGA and MIT cruises and various box sizes L ranging from 2 km to 64 km. The exponents are listed in Table 3.

Table 3. Comparison between theoretical and observed values of the exponent η.

Dataset	η_{obs}	η^0_{obs}	$\eta_{\text{th}} = 2\|\nu\|$
TOGA Cruise 1	0.760 ± 0.026	0.571	0.669 ± 0.025
MIT Cruise 1	0.728 ± 0.025	0.618	0.595 ± 0.047
TOGA Cruise 2	0.464 ± 0.011	0.382	0.477 ± 0.057
MIT Cruise 2	0.540 ± 0.014	0.417	0.409 ± 0.054
TOGA Cruise 3	0.582 ± 0.015	0.448	0.581 ± 0.056
MIT Cruise 3	0.573 ± 0.015	0.448	0.519 ± 0.063

Comparison between value of the power law index η defined by the stipulated behavior $p(L) \propto L^\eta$ of the fractional rainy area $p(L)$ as function of the averaging length scale L determined from the data and the value predicted by the model. The second column gives the "mean" exponent obtained by an overall least squares fit, while the third column gives the exponent determined from the slope at small L. In the last column specifying the "theoretical" value $\eta_{\text{th}} = 2|\nu|$ determined from the $\sigma^2(L)$ vs. L plots [Figure 2 via the fit to (24)], we also include the error bars not reported in Table 1.

formulas

$$p(L) = \begin{cases} 0.122(L/4)^{0.61} & \text{Phase I} \\ 0.098(L/4)^{0.68} & \text{Phase II} \end{cases},$$

which are consistent with the expected behavior $p(L) \to 0$ as $L \to 0$. [See also *Ha et al.*, 2002, who use a linear regression fit to describe $p(L)$ as function of L]. The exponents are, however, in each case quite different from the values suggested by the spectral model fit. We also find that for GATE, the quantity $\mu(L)$ actually depends significantly on L. In fact, $\mu(L)$ changes from about 1.77 at $L = 4$ km to about 2.16 at $L = 56$ km for Phase I and from about 1.89 at $L = 4$ km to about 2.22 at $L = 56$ km for Phase II.

Like in the TOGA COARE case, the goodness of a simple power law fit is probably somewhat fortuitous for GATE as well, since deviations from a power law behavior is to be expected on the basis of the model, because of the presence of "subdominant" terms in $1/p(L)$ such as the constant term in (43) or other possible L-dependent terms arising from a weak L-dependence of $\mu(L)$. Upon attempting to fit the data for $p(L)$ to a form like (43) we find that the best fit values of the exponent η are slightly smaller than the values given above. However, they still differ substantially from the model-inspired value $2|\nu|$. We have been unable to fully reconcile the apparent discrepancy between the two sets of exponents along the lines discussed above in connection with the TOGA COARE data.

One particular aspect that we did not address in this paper is the dependence of $p(L)$ and other rain statistics on the value of the rain threshold, i.e., the smallest value of rain rate that is experimentally measurable. This is an important issue, since different instruments have different thresholds and hence threshold dependence can greatly complicate intercomparison efforts for rain rate averages, especially conditional averages.

4.4. Discussion

From Table 1 it is evident that the model parameters vary considerably among the data sets. These variations are almost certainly due in part to real secular variations in the rainfall pattern. As discussed by *Short et al.* [1997], the TOGA COARE region of the tropical Western Pacific experiences planetary scale equatorial waves including an annual cycle, intraseasonal oscillations such as Madden-Julian oscillations [*Lau and Chan*, 1985; *Madden*, 1986; *Nakazawa* , 1995] and westerly wind bursts related to El Niño-Southern Oscillation (ENSO) events. The first half of Cruise 2 had an intensely active period associated with intraseasonal oscillation accompanied by passage of cloud superclusters and heavy rainfall. The rainfall time series during this period also shows a quasi 2-day oscillation of the type described by *Takayabu* [1994] and *Takayabu et al.* [1996]. They attribute this oscillation to westward-propagating inertio-gravity waves. By contrast, the spatial rainfall pattern during much of Cruise 1 appears to be much more homogeneous and isotropic, characteristic of small-scale convective systems. For a more detailed discussion see *Short et al.* [1997].

5. Conclusion

In this paper we have presented a model of second moment statistics of rainfall based on a linear stochastic dynamical equation. The model makes definite predictions about the statistical behavior of area-averaged rain rate in both space and time. In particular, the model predicts simple power law dependence of the rain statistics on the averaging area as the area tends to zero, perhaps reflecting a fractal nature of the rain field at smaller spatial scales. These predictions are tested with a radar-derived gridded tropical oceanic precipitation data set and the model is shown to describe the data quite well within the intrinsic limitations of the model assumption of statistical homogeneity, isotropy and stationarity. In particular, the intimate relationship between the dependence of the correlation length and time scales of the rain field on the spatial smoothing predicted by the model is in accordance with data. It should, however, be emphasized that in extrapolating the model-derived limiting behavior of the rain statistics deduced from a data set on a finite grid as the averaging scale goes to zero, we are effectively probing the "subgrid" structure of the rain field, i.e., the structure on scale finer than allowed by the instrumental resolution.

To conclude, we briefly mention some directions for future work. Like the spatially averaged rain rate, the spectral model also predicts relationships among the power law scaling exponents for the *time-averaged* rain rate as well. It would be interesting to see how well the model is capable of describing statistics of rain gauge data averaged over various time intervals. In a sequel to this paper, we will investigate the statistical properties of time-averaged local rain rates as measured by rain gauges using our model spectrum. Our model provides a convenient framework for addressing intercomparison problems that involve averaging the precipitation field over many different length and time scales. Recently the model has been used to investigate the problem of validation of satellite measurements of rainfall using rain gauge measurements on the ground [*Bell and Kundu*, 2002]. Finally, simple power law behavior of the various statistics of rain fields with respect to rescaling of the spatial averaging grid at small scales suggests a model-independent explanation based entirely on the scaling properties of the various statistical quantities involved. We plan to explore this possibility in future publications.

Appendix A: A Closed Analytic Formula for $\tau_{\rm int}(L)$

In this appendix we give a derivation of the formula (33) for the integral correlation time $\tau_{\rm int}(L)$ and describe its limiting behavior as $L \to 0$.

Equation (20) for the variance of the area-averaged rain rate in a grid box $\mathcal{A}$ of area $A = L^2$ gives the expression

$$\sigma_A^2 = 4\gamma_0 G(\nu; L/L_0) \,,$$

where $G(\nu; z)$ is defined as the double integral (21). On the other hand, the lagged autocovariance function $C_{\mathcal{A}\mathcal{A}}(0,\tau)$ of rain rate averaged over the same area is given by (BK96, Appendix A)

$$\begin{aligned} C_{\mathcal{A}\mathcal{A}}(0,\tau) &= \frac{2}{\pi}\int_0^\infty\int_0^\infty dk_1 dk_2\, \mathcal{G}^2\left(\frac{k_1 L}{2}\right) \\ &\quad \times \mathcal{G}^2\left(\frac{k_2 L}{2}\right) c(\mathbf{k},\tau) \\ &= \sqrt{\frac{2}{\pi}} F_0 \int_0^\infty\int_0^\infty dk_1 dk_2\, \mathcal{G}^2\left(\frac{k_1 L}{2}\right) \\ &\quad \times \mathcal{G}^2\left(\frac{k_2 L}{2}\right) \tau_k e^{-|\tau|/\tau_k} \end{aligned}$$

where $\mathcal{G}(x) = \sin^2 x/x^2$ is the Bartlett filter function. Setting $\tau = 0$, we have an alternative integral representation of σ_A^2 :

$$\begin{aligned} \sigma_A^2 &= \frac{2}{\pi}\int_0^\infty\int_0^\infty dk_1 dk_2\, \mathcal{G}^2\left(\frac{k_1 L}{2}\right)\mathcal{G}^2\left(\frac{k_2 L}{2}\right)\tau_k \\ &= \sqrt{\frac{2}{\pi}} F_0 \int_0^\infty\int_0^\infty \frac{dk_1 dk_2\, \mathcal{G}^2(k_1 L/2)\,\mathcal{G}^2(k_2 L/2)}{[1+(k_1^2+k_2^2)L_0^2]^{1+\nu}} \,, \end{aligned} \tag{A1}$$

where in the last step we have made use of the explicit form of τ_k given by (8). Comparing the two representations of σ_A^2 we obtain the useful integral identity

$$\begin{aligned} &\int_0^\infty\int_0^\infty \frac{dk_1 dk_2\, \mathcal{G}^2(k_1 L/2)\,\mathcal{G}^2(k_2 L/2)}{[1+(k_1^2+k_2^2)L_0^2]^{1+\nu}} \\ &= \frac{2\pi}{\Gamma(1+\nu)L_0^2} G(\nu; L/L_0) \,. \end{aligned} \tag{A2}$$

Next, consider the formula for the Fourier transform of $C_{\mathcal{A}\mathcal{A}}(0,\tau)$:

$$\begin{aligned} \tilde{C}_{\mathcal{A}}(\omega) &= \frac{2}{\pi}\int_0^\infty\int_0^\infty dk_1 dk_2\, \mathcal{G}^2\left(\frac{k_1 L}{2}\right) \\ &\quad \times \mathcal{G}^2\left(\frac{k_2 L}{2}\right) \hat{c}(k,\omega) \,. \end{aligned}$$

Using the explicit form of the model spectrum (9) with $\omega = 0$, we get

$$\tilde{C_{\mathcal{A}}}(\omega = 0) = \frac{2}{\pi} F_0 \int_0^\infty \int_0^\infty dk_1 dk_2 \, \mathcal{G}^2\left(\frac{k_1 L}{2}\right) \times \mathcal{G}^2\left(\frac{k_2 L}{2}\right) \tau_k^2$$

$$= \sqrt{\frac{2}{\pi}} F_0 \tau_0^2 \int_0^\infty \int_0^\infty \frac{dk_1 dk_2 \, \mathcal{G}^2(k_1 L/2) \, \mathcal{G}^2(k_2 L/2)}{[1 + (k_1^2 + k_2^2) L_0^2]^{2(1+\nu)}} .$$

In view of the identity (A2) this reduces to

$$\tilde{C_{\mathcal{A}}}(\omega = 0) = 4\sqrt{\frac{2}{\pi}} \gamma_0 \tau_0 \frac{\Gamma(1+\nu)}{\Gamma(2+2\nu)} G(1+2\nu; L/L_0) , \tag{A3}$$

where we have also used (14) to substitute for F_0 in terms of the standard model parameters γ_0, L_0 and ν. Using Eqs. (20), (A3) and (32) we obtain Eq. (33). The integrals were numerically evaluated with the help of *Mathematica* [v. 4; see *Wolfram*, 1999].

Finally, we examine the limiting behavior of $\tau_{\rm int}(L)$ as $L \to 0$. We note that for the range of values of the exponent ν of interest to us, $-1/2 < \nu < 0$, (or $0 < 1+2\nu < 1$), Eq. (22) yields the asymptotic behavior

$$G(\nu; z) = \frac{1}{2}\Gamma(|\nu|)\left(\frac{z}{2}\right)^{-2|\nu|} + O(1)$$

and

$$G(1+2\nu; z) = \frac{1}{8}\Gamma(1+2\nu) + O\left(z^{2(1+2\nu)}\right) ,$$

as $z \to 0$. From Eq. (33) it follows that $\tau_{\rm int}(L)$ vanishes for $L \to 0$ as

$$\tau_{\rm int}(L) \sim \frac{\tau_0}{4} \frac{\Gamma(1+\nu)}{(1+2\nu)\Gamma(|\nu|)} \left(\frac{L}{2L_0}\right)^{2|\nu|} . \tag{A4}$$

In this limit the product $\tau_{\rm int}(L)\sigma^2(L)$ approaches a constant :

$$\lim_{L\to 0} \tau_{\rm int}(L)\sigma^2(L) \simeq \frac{\gamma_0 \tau_0}{2} \frac{\Gamma(1+\nu)}{1+2\nu} . \tag{A5}$$

Appendix B: Effect of Finite Cut-off in the Evaluation of the Integral Correlation Time

In this appendix we obtain an expression for the fractional error in the integral correlation time $\hat{\tau}_{\rm int}$ predicted by the spectral model for a circular disk of radius a as a function of the lag $\tau_{\rm max}$ at which numerical integration of the correlation function is stopped. The derivation is a simple extension of the calculations described in *Bell and Kundu* [2002].

The covariance of instantaneous area-averaged rain rate averaged over a circular disk of radius a separated by a time interval τ is

$$\begin{aligned} C_{\mathcal{A}\mathcal{A}}(0,\tau) &\equiv \langle R'_A(\tau) R'_A(0) \rangle \\ &= \frac{1}{A^2} \int_A d^2\mathbf{x} \int_A d^2\mathbf{y} \, \langle R'(\mathbf{x},\tau) R'(\mathbf{y},0) \rangle \\ &= \frac{1}{A^2} \int_A d^2\mathbf{x} \int_A d^2\mathbf{y} \, c(\mathbf{x}-\mathbf{y},\tau) \\ &= \frac{1}{A^2} \int_A d^2\mathbf{x} \int_A d^2\mathbf{y} \, (2\pi)^{-3/2} \\ &\quad \times \int d^2\mathbf{k} \int d\omega \, e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} \tilde{c}(\mathbf{k},\omega) . \end{aligned}$$

Since $\tilde{c}(\mathbf{k},\omega)$ does not depend on the direction of $\mathbf{k}$, the areal integrals can be done using the identity

$$\begin{aligned} \int_A d^2\mathbf{x} \, e^{i\mathbf{k}\cdot\mathbf{x}} &= \int_0^a r dr \int_0^{2\pi} d\phi \, e^{ikr\cos\phi} \\ &= 2\pi a^2 J_1(ka)/(ka) , \end{aligned}$$

where $J_1(x)$ is the Bessel function of the first kind of order 1. The integral over ω with $\tilde{c}(\mathbf{k},\omega)$ as in Eq. (9) yields

$$\int_{-\infty}^{\infty} d\omega \, \tilde{c}(k,\omega) = \sqrt{2\pi} c(k,0) = \pi F_0 \tau_k .$$

After some algebra, one obtains

$$C_{\mathcal{A}\mathcal{A}}(0,\tau) = \frac{4\gamma_0 \Gamma(1+\nu)}{\alpha^2} \int_0^\infty \frac{d\kappa}{\kappa} \frac{J_1^2(\kappa\alpha)}{v(\kappa)} \times e^{-|\tau/\tau_0| v(\kappa)} . \tag{B1}$$

with the definitions

$$\alpha = a/L_0 , \tag{B2}$$

and

$$v(\kappa) = (1+\kappa^2)^{1+\nu} . \tag{B3}$$

For a circular disk of area L^2 the radius is $a = L/\sqrt{\pi}$.

Using (B1) one can easily show that

$$\tau_{\rm int} = \frac{\int_0^\infty d\tau \, C_{\mathcal{A}\mathcal{A}}(0,\tau)}{\sigma_A^2} = \tau_0 \left[\frac{\int_0^\infty \frac{d\kappa}{\kappa} \frac{J_1^2(\kappa\alpha)}{v^2(\kappa)}}{\int_0^\infty \frac{d\kappa}{\kappa} \frac{J_1^2(\kappa\alpha)}{v(\kappa)}} \right] . \tag{B4}$$

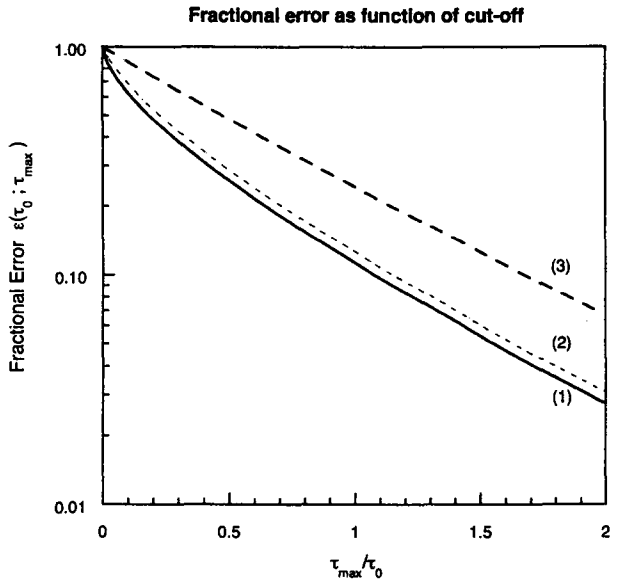

Figure B1. Plot of the fractional error as function of the ratio $\tau_{\max}/\tau_0$ computed for the typical spectral model parameters $\nu = -0.25$ and $L_0 = 70$ km and three values of disk radius a. The curves marked (1), (2) and (3) represent the results for the radii $a = 1$ km, 10 km and 100 km. The corresponding true integral correlation times are $0.052\tau_0$, $0.19\tau_0$ and $0.65\tau_0$ respectively.

If we estimate $\tau_{\rm int}$ by introducing a finite upper cut-off in the τ-integral, then the fractional error $\epsilon(\tau_0;\tau_{\max})$ defined as

$$\epsilon(\tau_0;\tau_{\max}) = \frac{\tau_{\rm int} - \hat{\tau}_{\rm int}}{\tau_{\rm int}} \ . \tag{B5}$$

is given by the formula

$$\epsilon(\tau_0;\tau_{\max}) = \frac{\int_0^\infty \frac{d\kappa}{\kappa}\frac{J_1^2(\kappa\alpha)}{v^2(\kappa)}\exp\left[-(\tau_{\max}/\tau_0)v(\kappa)\right]}{\int_0^\infty \frac{d\kappa}{\kappa}\frac{J_1^2(\kappa\alpha)}{v^2(\kappa)}} \ . \tag{B6}$$

Eq. (B6) is numerically evaluated using *Mathematica* [*Wolfram*, 1999] to estimate the correction to the integral correlation time discussed in section 3.3. A plot of ϵ against the ratio $\tau_{\max}/\tau_0$ is shown in Figure B1 for typical parameter values.

Appendix C: Estimate of Random Error for the Integral Correlation Time of an AR(1) Process

In this appendix we obtain a simple estimate of the random error of the integral correlation time of a normally distributed random variable undergoing a stationary, linear, first order autoregressive [AR(1)] process, evaluated from a single time series of finite length T.

Consider the time series $\{x(t)\}$ of length T generated by a single realization of a random variable $X(t)$ undergoing a linear stationary AR(1) process for which the lagged covariance function

$$\gamma(\tau) = \langle X'(t+\tau)X'(t)\rangle \tag{C1}$$

has the simple exponential form

$$\gamma(\tau) = \sigma^2 e^{-|\tau|/\tau_A} \ . \tag{C2}$$

Its estimate, obtained from a finite time series, namely,

$$c(\tau) = \frac{1}{T}\int_0^T dt x'(t+\tau)x'(t)$$

can be thought of as a random variable

$$Y(\tau) = \frac{1}{T}\int_0^T dt X'(t+\tau)X'(t) \tag{C3}$$

depending on the realization of $X(t)$. We are interested in estimating the random error for the integral correlation time represented by the random variable

$$T_{\rm int} = \int_0^{\tau_{\max}} Z(\tau)d\tau \tag{C4}$$

where the variable $Z(\tau) = Y(\tau)/Y(0)$ represents the lagged correlation function. The desired error estimate is simply the square root of the variance

$$\begin{aligned}\sigma^2[T_{\rm int}] &= \langle (T_{\rm int} - \langle T_{\rm int}\rangle)^2\rangle \ , \\ &= \langle T_{\rm int}^2\rangle - \langle T_{\rm int}\rangle^2 .\end{aligned}$$

The variance $\sigma^2[T_{\rm int}]$ can then be expressed in the form

$$\sigma^2[T_{\rm int}] = \int_0^{\tau_{\max}}\int_0^{\tau_{\max}} d\tau_1 d\tau_2 \langle Z'(\tau_1)Z'(\tau_2)\rangle \ . \tag{C5}$$

For simplicity of notation we write

$$Y(\tau) = A, Y(0) = B.$$

These random variables can each be expressed as the sum of respective mean values and fluctuations:

$$A = a + \alpha, \ B = b + \beta,$$

where

$$a \equiv \langle A\rangle = \gamma(\tau) \ , \ b \equiv \langle B\rangle = \sigma^2 \ .$$

Since up to second order in the fluctuations

$$Z(\tau) = \frac{A}{B} \approx \frac{a}{b}\left(1+\frac{\alpha}{a}\right)\left(1-\frac{\beta}{b}+\frac{\beta^2}{b^2}\right) ,$$

it follows that

$$\langle Z(\tau)\rangle = \left\langle \frac{A}{B}\right\rangle \approx \frac{a}{b}\left(1+\frac{\langle\beta^2\rangle}{b^2}-\frac{\langle\alpha\beta\rangle}{ab}\right) ,$$

where the cubic and higher order terms are neglected. In this approximation

$$\begin{aligned}\langle Z'(\tau_1)Z'(\tau_2)\rangle &= \left\langle \frac{A_1}{B}\cdot\frac{A_2}{B}\right\rangle - \left\langle\frac{A_1}{B}\right\rangle\left\langle\frac{A_2}{B}\right\rangle \\ &\approx \frac{1}{b^2}\left\langle \alpha_1\alpha_2+\frac{\beta^2}{b^2}a_1a_2\right\rangle , \qquad \text{(C6)}\end{aligned}$$

where the subscripts 1 and 2 indicate evaluation at lags τ_1 and τ_2 respectively.

Eq. (C6) can be recast in a much simpler form if $X(t)$ (and therefore $X'(t)$) is a gaussian random variable. Using the relation [see, e.g., *Anderson*, 1958]

$$\begin{aligned}\langle X'(t_1)X'(t_2)X'(t_3)X'(t_4)\rangle = \\ \langle X'(t_1)X'(t_2)\rangle\langle X'(t_3)X'(t_4)\rangle \\ + \langle X'(t_1)X'(t_3)\rangle\langle X'(t_2)X'(t_4)\rangle \\ + \langle X'(t_1)X'(t_4)\rangle\langle X'(t_2)X'(t_3)\rangle ,\end{aligned}$$

expressing the fourth moment of such a variable in terms of products of the second moment, the quantity

$$\begin{aligned}\langle\alpha_1\alpha_2\rangle &= \langle Y'(\tau_1)Y'(\tau_2)\rangle \\ &= \langle Y(\tau_1)Y(\tau_2)\rangle - \langle Y(\tau_1)\rangle\langle Y(\tau_2)\rangle\end{aligned}$$

can be reduced to the form

$$\begin{aligned}\langle\alpha_1\alpha_2\rangle = \frac{1}{T^2}\int_0^T\int_0^T dt_1dt_2\,[&\gamma(t_1-t_2+\tau_1-\tau_2)\gamma(t_1-t_2) \\ &+\gamma(t_1-t_2+\tau_1)\gamma(t_1-t_2-\tau_2)] . \qquad \text{(C7)}\end{aligned}$$

The double integrals can be easily performed with the help of the integral identity

$$\int_0^T\int_0^T dt_1dt_2\, f(t_1-t_2) = \int_{-T}^T d\tau\,(T-|\tau|)f(\tau) , \qquad \text{(C8)}$$

which, for a function $f(\tau)$ decaying exponentially as $\tau\to\infty$, in the approximation $T \gg \tau_A$ further simplifies to

$$\int_0^T\int_0^T dt_1dt_2\, f(t_1-t_2) \approx T\int_{-\infty}^{\infty} d\tau'\, f(\tau') .$$

Eq. (C7) then reduces to

$$\begin{aligned}\langle\alpha_1\alpha_2\rangle &\approx \frac{1}{T}\int_{-\infty}^{\infty} d\tau'\,[\gamma(\tau'+\tau_1-\tau_2)\gamma(\tau') \\ &\quad +\gamma(\tau'+\tau_1)\gamma(\tau'-\tau_2)]. \qquad \text{(C9)}\end{aligned}$$

Specializing Eq. (C9) to the case $\tau_1=\tau_2=0$, we also have

$$\langle\beta^2\rangle \approx \frac{2}{T}\int_{-\infty}^{\infty} d\tau'\,\gamma^2(\tau'). \qquad \text{(C10)}$$

Making use of the explicit form of the lagged covariance function $\gamma(\tau)$ given by Eq. (C2) and carrying out the remaining integration, we get

$$\begin{aligned}\langle\alpha_1\alpha_2\rangle &\approx \frac{\sigma^4}{T}\left[(|\tau_1-\tau_2|+\tau_A)e^{-|\tau_1-\tau_2|/\tau_A}\right. \\ &\quad \left.+(\tau_1+\tau_2+\tau_A)e^{-(\tau_1+\tau_2)/\tau_A}\right] , \qquad \text{(C11)}\end{aligned}$$

$$\langle\beta^2\rangle \approx \frac{2\sigma^4\tau_A}{T} . \qquad \text{(C12)}$$

Eq. (C5) then becomes

$$\begin{aligned}\sigma^2[T_{\rm int}] &\approx \frac{1}{T}\int_0^{\tau_{\max}}\int_0^{\tau_{\max}} d\tau_1d\tau_2 \\ &\quad\times\left[(|\tau_1-\tau_2|+\tau_A)e^{-|\tau_1-\tau_2|/\tau_A}\right. \\ &\quad\left.+(\tau_1+\tau_2+\tau_A)e^{-(\tau_1+\tau_2)/\tau_A}\right] \\ &\quad+\frac{2\tau_A}{T}\left[\int_0^{\tau_{\max}} d\tau' e^{-|\tau'|/\tau_A}\right]^2 \qquad \text{(C13)}\end{aligned}$$

The double integrals in the first term can be simplified using Eq. (C8) with $\tau_{\max}$ instead of T. In the approximation $\tau_{\max}\gg\tau_0$, we finally obtain the desired formula

$$\sigma^2[T_{\rm int}] \approx \frac{4\tau_{\max}\tau_A^2}{T}\left[1+O\left(\tau_A/\tau_{\max}\right)\right] , \qquad \text{(C14)}$$

which is Eq. (42), since $\tau_0=\tau_A$ for the exponentially decaying correlation (C2).

Acknowledgments. This research was supported by the Office of Earth Science of the National Aeronautics and Space Administration as part of the Tropical Rainfall Measuring Mission.

References

Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, 1958.

Bell, T. L., Climate sensitivity from Fluctuation Dissipation: Some simple model tests, *J. Atmos. Sci.*, *37*, 1700–1707, 1980.

Bell, T. L., A space–time stochastic model of rainfall for satellite remote-sensing studies, *J. Geophys. Res.*, *92D*, 9631–9643, 1987.

Bell, T. L., A. Abdullah, R. L. Martin and G. R. North, Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space–time stochastic model, *J. Geophys. Res.*, *95D*, 2195–2205, 1990.

Bell, T. L., and N. Reid, Detecting the diurnal cycle of rainfall using satellite observations, *J. Appl. Meteor.*, *32*, 311–322, 1993.

Bell, T. L., and P. K. Kundu, A study of the sampling error in satellite rainfall estimates using optimal averaging of data and a stochastic model, *J. Climate*, *9*, 1251–1268, 1996.

Bell, T. L., P. K. Kundu and C. D. Kummerow, Sampling errors of SSM/I and TRMM rainfall averages: Comparison with error estimates from surface data and a simple model, *J. Appl. Meteor.*, *40*, 938–954, 2001.

Bell, T. L., and P. K. Kundu, Comparing satellite rainfall estimates with rain-gauge data: Optimal strategies suggested by a spectral model, *J. Geophys. Res.*, in press, 2002.

Cox, D. R., and V. Isham, A simple spatial-temporal model of rainfall, *Proc. R. Soc. Lond.*, A *415*, 317–328, 1988.

Gupta, V. K., and E. C. Waymire, A stochastic kinematic study of subsynoptic space-time rainfall, *Water Resour. Res.*, *15*, 637–644, 1979.

Gupta, V. K., and E. C. Waymire, Multiscaling properties of spatial rainfall and river flow distributions, *J. Geophys. Res.*, *95D*, 1999–2010, 1990.

Ha, E., G. R. North, C. Yoo and K.-J. Ha, Evaluation of some ground truth designs for satellite estimates of rain rate, *J. Atmos. Oceanic Tech.*, *19*, 65–73, 2002.

Islam, S., R. L. Bras, and I. Rodríguez-Iturbe, Multidimensional modeling of cumulative rainfall: Parameter estimation and model adequacy through a continuum of scales, *Water Resour. Res.*, *24*, 992–995, 1988.

Jenkins, G. M., and D. G. Watts, *Spectral Analysis and its applications*, Holden-Day, San Francisco, 1968.

Kedem, B., L. S. Chiu and G. R. North, Estimation of rain rate: Application to satellite observations, *J. Geophys. Res.*, *95D*, 1965–1972, 1990.

Koepsell, R. W., and J. B. Valdés, Sampling properties of parameter estimators for a storm field rainfall model, *Water Resour. Res.*, *25*, 2067–2085, 1989.

Krajewski, W. F., and J. A. Smith, Sampling properties of parameter estimators for a storm field rainfall model, *Water Resour. Res.*, *25*, 2067–2085, 1989.

Lau, K.-M. and P. H. Chan, Aspects of the 40–50-day oscillation in the tropics, *Mon. Wea. Rev.*, *113*, 1889–1909, 1985.

Laughlin, C. R., On the effect of temporal sampling on the observation of mean rainfall, in *Precipitation Measurements from Space, Workshop Report*, edited by D. Atlas and O. W. Thiele, pp. D59-D66, available from Goddard Space Flight Center, Greenbelt, MD 20771, 1981.

Le Cam, L., A stochastic description of precipitation, in *IV Berkeley Symposium in Mathematical Statistics and Probability*, vol. 3, pp. 165–186, University of California, Berkeley, 1961.

Lovejoy, S. Area-perimeter relations for rain and cloud areas, *Science*, *187*, 1035–1037, 1982.

Lovejoy, S. and B. Mandelbrot, Fractal properties of rain and a fractal model, *Tellus*, *37A*, 209, 1985.

Madden, R. A., Seasonal variations of the 40–50-day oscillation in the tropics, *J. Atmos. Sci.*, *43*, 3138–3158, 1986.

McConnell, T. L., and G. R. North, Sampling errors in satellite estimates of tropical rainfall: Monte Carlo study using a space–time stochastic model, *J. Geophys. Res.*, *92D*, 9567–9570, 1987.

Nakazawa, T., Intraseasonal oscillations during the TOGA COARE IOP, *Proc. Meteor. Soc. Japan*, *73*, 305–319, 1995.

North, G. R., and S. Nakamoto, Formalism for comparing rain estimation designs, *J. Atmos. Oceanic Tech.*, *6*, 985–992, 1989.

Rodríguez-Iturbe, I., D. R. Cox, and V. Isham, Some models for rainfall based on stochastic point processes, *Proc. Royal Soc. Lond.*, *A 410*, 269–288, 1987.

Schertzer, D., and S. Lovejoy, Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes, *J. Geophys. Res.*, *92D*, 9693–9714, 1987.

Short, D. A., D. B. Wolff, D. Rosenfeld, and D. Atlas, A study of the threshold method utilizing raingage data, *J. Appl. Meteorol.*, *32*, 1379–1387, 1993.

Short, D. A., P. A. Kucera, B. S. Ferrier, J. A. Gerlach, S. A. Rutledge, and O. W. Thiele, Shipborne radar rainfall patterns within the TOGA COARE IFA, *Bull. Amer. Meteorol. Soc.*, *78*, 2817–2836, 1997.

Smith, J. A., and A. F. Karr, Parameter estimation for a model of space-time rainfall, *Water Resour. Res.*, *21*, 1251–1257, 1985.

Smith, J. A., and W. F. Krajewski, Statistical modeling of space-time rainfall using radar and rain gage observations, *Water Resour. Res.*, *23*, 1893–1900, 1987.

Takayabu, Y. N., Large-scale cloud disturbances associated with equatorial waves. Part II: Westward propagating inertio-gravity waves, *Proc. Meteor. Soc. Japan*, *72*, 451–465, 1994.

Takayabu, Y. N., K.-M. Lau, and C.-H. Sui, Two-day variation in Cloud-atmosphere coupled system observed in TOGA COARE IOP, *Mon. Wea. Rev.*, *124*, 1892–1913, 1996.

Valdés, J. B., I. Rodríguez-Iturbe, and V. K. Gupta, Approximations of temporal rainfall from a multidimensional model, *Water Resour. Res.*, *21*, 1259–1270, 1985.

Valdés, J. B., S. Nakamoto, S. S. P. Shen, and G. R. North, Estimation of multidimensional precipitation parameters by areal estimates of oceanic rainfall, *J. Geophys. Res.*, *95 D*, 2101–2111, 1990.

Valdés, J. B., E. Ha, C. Yoo, and G. R. North, Stochastic characterization of space-time precipitation: Implications for remote sensing, *Adv. Water Resour.*, *17*, 47–59, 1994.

Waymire, E., V. K. Gupta, and I. Rodríguez-Iturbe, A spectral theory of rainfall intensity at the meso-β scale, *Water Resour. Res.*, *20*, 1453–1465, 1984.

Waymire, E., Scaling limits and self-similarity in precipitation fields, *Water Resour. Res.*, *21*, 1271–1281, 1985.

Webster, P. J., and R. Lukas, TOGA COARE: The Coupled Ocean-Atmosphere Experiment, *Bull. Amer. Meteorol. Soc.*, *73*, 1377–1416, 1992.

Wheater, H. S., V. S. Isham, D. R. Cox, R. E. Chandler, A. Kakou, P. J. Northrop, L. Oh, C. Onof, and I. Rodríguez-Iturbe, Spatial-temporal rainfall fields: modelling and statistical aspects, *Hydrol. Earth System Sci.*, *4*, 581–601, 2000.

Wolfram, S., *The Mathematica Book*, 4th ed., Wolfram Media/Cambridge University Press, 1999.

Yoo, C., J. B. Valdés, and G. R. North, Stochastic modeling of multidimensional precipitation fields considering spectral structure, *Water Resour. Res.*, *32*, 2175–2187, 1996.

Prasun K. Kundu and Thomas L. Bell, NASA / Goddard Space Flight Center, Mail Code 913, Greenbelt, MD 20771, U.S.A. (kundu@climate.gsfc.nasa.gov; bell@climate.gsfc.nasa.gov)

This preprint was prepared with AGU's LaTeX macros v5.01, with the extension package 'AGU++' by P. W. Daly, version 1.6b from 1999/08/19.

within about 1% of what was reported in BK96 from a direct fit to the lagged autocorrelation function.

4. Results and Discussion

The full set of parameter values for which the model best fits the TOGA COARE data are given in Table 1, along with the corresponding values for GATE Phase I determined in BK96 using a somewhat less exact (and more arduous) trial-and-error curve fitting method instead of the analytical method adopted here.

4.1. Model Parameters ν, L_0 and γ_0

Table 1 summarizes the TOGA COARE model parameters. We see that the values of the power law exponent ν range between about -0.21 to -0.34 indicating a much stronger singular behavior than was found for the GATE Phase I data ($\nu = -0.11$). The characteristic time scales τ_0 lie in the range of roughly 4.5 to 8.2 hours, considerably shorter than the GATE Phase I value $\tau_0 \approx 13$ hours. The characteristic length L_0 ranges between about 54 and 94 km indicating a somewhat sharper fall-off of the spatial correlations compared to GATE Phase I for which $L_0 = 104$ km. The results exhibiting the quality of the fit are shown in Figure 2 for the TOGA and MIT radars. It is seen that in most cases the fit is quite good up to grid size $L = 64$ km, beyond which the approximation on which Eq. (24) is based is expected to break down. Our results seem to be consistent with the model prediction that the variance $\sigma^2(L)$ has a power law behavior $\sigma^2(L) \sim L^{-2|\nu|}$ as $L \to 0$. Robustness of the parameterization was checked by comparing the variance of area-averaged rain rate $(\sigma_A^2)_{\rm th}$ for an area $A = L^2$ with $L = 128$ km predicted by the exact formula (20), rather than the asymptotic formula (24), against the observed variance $(\sigma_A^2)_{\rm obs}$ for the 128 km box centered at the two ship locations. Table 2 shows the comparisons, which are quite good.

For GATE data we find that for L between 4 km and 56 km, the variance $\sigma^2(L)$ is quite accurately fitted (not shown here) by the formulas

$$\sigma^2(L) = \begin{cases} 17.3L^{-0.24} - 4.66 \text{ mm}^2\text{h}^{-2} & \text{Phase I} \\ 15.0L^{-0.38} - 1.92 \text{ mm}^2\text{h}^{-2} & \text{Phase II} \end{cases},$$

which through Eqs. (24) to (26) yield the parameter values

$$\gamma_0 = 1.02 \text{ mm}^2\text{h}^{-2},\ \nu = -0.12,\ L_0 = 93.8 \text{ km}$$

for Phase I and

$$\gamma_0 = 0.63 \text{ mm}^2\text{h}^{-2},\ \nu = -0.19,\ L_0 = 82.7 \text{ km}$$

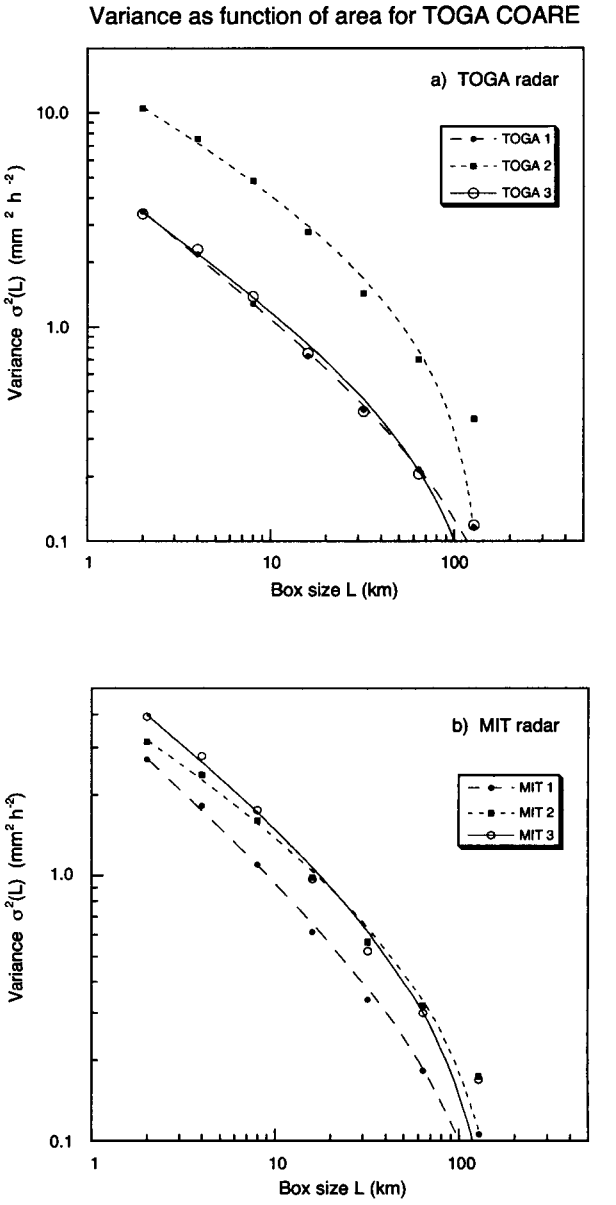

Figure 2. Comparison between the observed variance of area-averaged rain rate $\sigma^2(L)$ for the TOGA COARE data and the corresponding spectral model predictions for various box sizes L ranging from 2 km to 128 km using Eqs. (24)–(26) and the parameter values listed in Table 1.